Processing savings and work decisions through Shannon's channels

Tutino, Antonella

Federal Reserve Board

22 May 2008

Processing savings and work decisions through Shannon's channels

Antonella Tutino *

First version: May 2008 ; This version: July 2009

Abstract

This paper argues that constraining people to choose consumption and labor under finite Shannon capacity produces results in line with U.S. business cycle data as well as secular movements in consumption and labor supply. The model has a simple partial equilibrium setting in which risk averse consumers keep high labor supply and low consumption profile at early stage of life to hedge against wealth fluctuations. They rationally choose to keep consumption and labor unchanged until they collect enough information. I find that at high frequency consumption appears to be more sluggish than labor supply. However, when people decide to change consumption they do so by a large amount. This combination leads to higher variance of consumption with respect to labor supply. The model also finds high persistence and strong comovement of consumption and employment and delayed response of consumption and labor with respect to shocks to wages. Positive changes in wages generate an increase in long run value of consumption while the change in long run values of labor is negligible. Furthermore, the effects on labor and consumption of a shock to wages propagate slowly over time due to people's endogenous choice of information. These findings suggest that rational inattention offers a promising avenue to bridge the gap between theory and U.S. data at business cycle frequency as well as in the long run.

*E-mail: *Antonella.Tutino@frb.gov.* The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Federal Reserve Board or any other person associated with the Federal Reserve System.

1 Introduction

Existing macroeconomic theories have trouble fitting empirical regularities in consumption and labor at both high and low frequencies. Existing models match some U.S. business cycle facts at the expense of secular facts. Despite the progress in the field and rich modelling tools, one peculiar dimension of the business cycle that still troubles the literature is the labor market. As a matter of theory, business cycle and growth facts on consumption and labor should emerge from a framework in which consumers continuously solve dynamic optimization problems. As a matter of data, most of the dynamic optimization models proposed by modern macroeconomic theory generate predictions at odds with U.S. evidence.[1]

The trouble is that specifications of neoclassical rational expectations model that capture consumption and labor movements at business cycle frequencies seem inadequate to explain postwar movements in these variables. Starting from Lucas and Rapping (1969), several papers have tried to reconcile theory and facts by focussing on macro and micro estimates of labor supply elasticities,[2] identifying labor supply and demand through search and matching models[3] and questioning the validity of dynamic optimization framework as an appropriate representation of people's actual consumption and labor behavior.[4]

This paper suggests an explanation of U.S. business cycle facts and secular facts based on rational inattention.[5] My choice to model people constrained by finite information processing capacity agrees with intuition and, as the paper shows, with the empirical evidence.

In my model, rational households consume, work and pay attention to their wealth to maximize their lifetime utility. In the baseline model, I assume that consumers are risk adverse and have a constant Frisch elasticity of substitution. Under rational inattention, they cannot know the exact value of their wealth in each period due to information-processing frictions. Each period they choose information about wealth within the limits of their capacity and decide on the basis of that information how much to consume and work. Realized consumption and employment are used to update rationally their knowledge of wealth. In my setting wages follows a Markov process and its distribution is known before consumers make their work and consumption decisions. Thus, fluctuations in wealth are due to movements in labor and consumption as well as movements in wages. Consumers keep track of wealth by processing information through a Shannon channel.

Having a bound on information processing rate suits the observation that people do not check their account on a daily basis, nor they are likely to keep track of the incidence of their expenses and hours worked on their lifetime wealth at high frequency. Using

[1] See, e.g., Mankiw, Rotemberg and Summers (1985), Barro and King (1984), Hall (1997) and Chang and Kim (2005).

[2] See, e.g., King, Plosser and Rebelo (1988), Rogerson and Wallenius (2007) and Shimer (2005, 2008).

[3] See, e.g., Rotemberg (1998), Rotemberg and Woodford (1997), Shimer (2005).

[4] See, e.g., Mankiw, Rotemberg and Summers (1985).

[5] cfr., Shannon (1948), Sims (1988, 1998, 2003, 2006), Tutino (2008).

Shannon channels as modelling device has also the nice feature of providing a natural bound for information flow which depends only on the distribution of the variables that are passed through the channel, regardless of the characteristic of such a channel. This in turns makes the choice of this kind of information friction free from unexplained assumptions on individual characteristics.

More importantly, the predictions of the model agree with U.S. data along the dimensions analyzed. I show that even in a simple partial equilibrium setting with focus on the supply side of the labor market, a model in which people choose labor and consumption under information processing constraint à la Sims is able to explain several features of the data. In particular, my setting delivers four results: 1. the ratio of standard deviation of consumption over hours worked bigger than one; 2. persistence in consumption and labor supply; 3. comovement of consumption, labor and wealth; 4. endogenously persistent propagation of shocks to wages that delivers a long-term variation in consumption's growth and a negligible effect on long-term labor supply.

To understand the mechanism behind these results, consider what happens in an equivalent model with full information. With stochastic wages and interest rates consistent with consumption smoothing, a consumer with log utility and constant Frisch elasticity accumulates wealth during early periods of his life by limiting consumption and increasing work effort. In such a context, fluctuations in wealth are mostly due to fluctuations in labor income rather than consumption plans. With information processing constraints, consumers cannot know the exact value of their wealth. They keep track of their wealth imperfectly by choosing signals as informative about wealth as channel capacity allows them to. Log utility in consumption and convex disutility of labor together with low information flow make households work hard and save at early stage of their life. With low information flow and, as a result, low informativeness of the signal each period, households keep savings and labor supply high to make sure that they can sustain their consumption. As wealth accumulates, the signal on high values of wealth becomes sharper, calling for a major adjustment in behavior. The size of the adjustment is bigger the lower their processing capacity. This result is intuitive: the longer consumers wait to modify their behavior, the bigger is the variation in consumption and labor once they acknowledge the change in wealth through information processed. Furthermore, in my model, consumption is more sluggish than labor supply. The rationale for this finding lies in the preferences of consumers who dislike having to change their consumption frequently while having constant Frisch elasticity of labor supply. Such preferences imply that people review their consumption plan significantly when they realize that they have saved too much or too little with respect to their lifetime possibilities. This mechanism in turns leads to higher volatility of realized consumption with respect to labor supply and provides a rationale for the first result. The effect is stronger the lower the information flow.

The autocorrelation of consumption and employment stems from a similar logic. With signals that bring about low information, changes in behavior are slow at high frequency: the news about wealth are not enough to modify yesterday's consumption and labor supply. This mechanism implies that consumers maintain the same consumption and labor profile and keep accumulating information until evidence of changes in wealth suggests otherwise.

As for the third finding, my model predicts stronger comovement of contemporaneous consumption, labor and wealth the higher the information flow is. If information capacity is low, then contemporaneous consumption and labor comove strongly with lagged values of wealth. Contemporaneous consumption, labor and wealth are strongly linked via the budget constraint. Moreover, the budget constraint affects the choice of the policy function -i.e., optimal joint distribution between wealth and behavior-. With high information flow, consumers' optimal policy commands wealth and behavior to be as related as possible so that the outcome from consumption and labor supply are very informative about wealth. If information flow is low, consumption and labor are strongly correlated with past values of wealth. This result is driven by the rational (Bayesian) update of consumers' information. Each period the household gets information about wealth and observes consumption and labor choices. Low information flows makes the signal on wealth imprecise forcing the household to rely mostly on the information content of his consumption and labor. This translates into periods of inertial behavior until the information collected signals enough variation in wealth to justify a change in the choices. This observation together with Bayesian updating explain why, in presence of finite rate of information, consumers delay their reaction to changes in wealth. The results are robust to several utility specifications.

The last result of the model concerns the short and long term response of consumption and labor to fluctuations in wages. Following an increase in wage, in the short run both consumption and labor increase due to an income effect generated by the interplay of the curvature of the utility function with the information processing constraint. People slowly process signals on the increase in wealth due to savings and labor income. While processing information, consumption and employment behaviors exhibit a sluggish dynamic. When people realize that wealth has grown, they increase consumption permanently and start enjoying more leisure. Thus, in the model, the long term effect of an increase in wage leads to an higher long run value of consumption with respect to its pre-shock path while the long run value of employment is similar to the pre-shock case. These short and long run responses of consumption and labor following a change in wage as well as the transitional dynamics are consistent with the U.S. evidence documented in the real business cycle literature as well as U.S. growth's facts. [6]

Together with the contribution to the macroeconomic labor literature, this paper is closely related to the literature of rational inattention, with particular reference to Sims (2003, 2006), Tutino (2008) and Mackowiak and Wiederholt (2008a, 2008b). This paper departs from Sims (2003) and Mackowiak and Wiederholt (2008a, 2008b) since the ex-ante characterization of uncertainty is not limited to the Gaussian distribution nor the framework is constrained to be linear quadratic. Instead, as in Sims (2006) and Tutino (2008), this paper presents a fully endogenous choice of distribution of uncertainty in a dynamic context and allows for risk aversion in the specification of agents' preferences.

[6] See, e.g., Lucas and Rapping (1969), MaCurdy (1981), Hall , Cooley and Nason (1995).

The present framework extends Tutino (2008) by augmenting the choice space of people to consumption and labor as opposite to only consumption. This extension generates endogenously an allocation of attention between the two activities -consumption and work-, and, in turn, a different degree of persistence between consumption and labor on the basis of the stochastic properties of the joint distribution of actions and wealth chosen by the consumers. An example might help clarifying the intuition behind this property. Suppose that a person works the same amount of hours everyday. Given this behavior, the person learns nothing about wealth through his labor supply. In this case, fluctuations of wealth are acknowledged mostly through the information content of consumption realizations and the optimal distribution of uncertainty is similar to the one derived in Tutino (2008). The household might have a better understanding of his financial possibilities by varying either consumption or leisure or both and thereby improve his utility. If wages are relatively stable, it might be optimal to keep labor supply fixed and offset fluctuations in wealth with changes in consumption. On the other hand, if wages change significantly, it might pay off to vary the amount of labor supply and maintain a smooth consumption profile. Amount and directions of these changes in behavior depend on the relative cost of changing consumption with respect to labor supply, the relative benefits of being better informed about wealth through either source and households' preferences implied by the curvature of the utility function. Thus, in this framework, predominance of income vs. substitution effect does not depend only on people's utility as standard macro literature delivers, but also on the relative attention that people pay to current consumption and hour worked as source of information about wealth.

The rest of the paper is structured as follows: The next section presents the model and its main assumptions while section 3 illustrates the computation strategy. Section 4 is the core of the paper. It analyzes the main findings and contrasts them with US data. Section 4 also provides statistics and predictions of different specifications of the model. Section 5 shows the properties of the optimal policy whereas Section 6 concludes. Appendices cover the mathematical proofs (Appendix A), additional statistics for the prediction of the model (Appendix B), a low-dimension analytical solution for the model (Appendix C) and the pseudocode (Appendix D).

2 The Model

The model is a one sector partial equilibrium discrete time problem. To fix notation and intuition, first I discuss the model without information processing constraint. Then, I introduce information processing constraints à la Shannon and present the full rational inattention model.

2.1 A version of the model under infinite processing capacity

The economy is populated by numerous households who maximize the expected discounted value of their utility. Utility, $u(C,L)$, is defined over a consumption good, C, and labor, L and is strictly increasing and strictly concave in both its arguments, $\lim_{C\to 0} u_C(C,L) = +\infty, \forall L \in [0,1]$, $\lim_{L\to 0} u_L(C,L) = +\infty \ \forall C \geq 0$. Moreover, I assume that utility is homogeneous and additively separable in consumption and leisure. In particular:

$$u(C_t, L_t) = \log(C_t) - \frac{\alpha}{\eta+1} L_t^{\eta+1} \tag{1}$$

where η is the inverse of Frisch elasticity of labor supply, $\varepsilon > 0$, and $\alpha > 0$ is a constant disutility associated to labor.

Each period, people are endowed with one unit of time ($L \leq 1$) and face a stochastic real wage, s, in exchange for their labor effort. Wages follow a stationary i.i.d. Markov process with transition $\pi_s(s'|s) = \Pr(s_{t+1} \leq s'|s_t = s)$.

Consumers' wealth, W, evolves according to previous' period savings $(W - C)$, augmented with a fixed and exogenous interest rate, R, and labor income, $s * L$. Given the assumption on the wage process, the problem is stationary. The recursive formulation of the household's problem is:

$$V(W_t) = \max_{C_t, L_t} u(C_t, L_t) + \beta E_t V(W_{t+1}) \tag{2}$$

s.t.

$$W_{t+1} = R(W_t - C_t) + s_t L_t \tag{3}$$

$$W_0 \text{ given} \tag{4}$$

$$L_t \leq 1, \quad C_t \geq 0 \ \forall (L, C), \forall t \tag{5}$$

Note that in this setting the only source of uncertainty is wage, s. So long as wages are a Markov i.i.d. process -as assumed-, uncertainty about wages translates directly into uncertainty about next period's wealth. It follows that the initial condition on wealth, (4), is equivalent to knowing s_0. Expectations of the Bellman value next period - $V(W')$- are taken conditional on the current value of W. Moreover, I assume that $\beta R = 1$. Optimality conditions of the household with respect to consumption and labor imply the following contemporaneous relation

$$L_t = \left(\frac{s_t}{C_t \alpha}\right)^{\frac{1}{\eta}} \tag{6}$$

So long as both η and α are finite and with a desire to smooth consumption implied by (1), condition (6) implies that labor will change through time reflecting changes in wages.

The intertemporal optimal condition for consumption is:

$$1 = E_t\left(\frac{C_{t+1}}{C_t}\right) \tag{7}$$

and the equivalent for labor

$$\frac{L_t^{\eta}}{s_t} = E_t\left(\frac{L_{t+1}^{\eta}}{s_{t+1}}\right) \tag{8}$$

To match the joint behavior of per capita consumption and per capita hours that we observe in US data, the model should produce (1) cyclical movements: procyclical behavior of per capita consumption and per capita hours worked; (2) secular movements: labor supply's response to permanent changes in wages is negligible while consumption's responses are significant;

As Mankiw, Rotemberg and Summers (1985) noted, the first order conditions (6)-(8) cannot simultaneously account for facts (1) and (2). They prove that U.S. data strongly reject specifications of the kind (6)-(8), questioning the validity of continuous dynamic optimization as useful framework to match data on consumption and labor. The next section proposes a dynamic optimization problem based on rational inattention theory whose predictions are in line with observed time series of consumption and labor.

2.2 Rational inattention version of the model

Under *Rational Inattention Theory* (Sims, 2003, 2006), information is fully and freely available to the agents. However, people cannot process quickly and precisely all the information due to Shannon's processing constraints. Recognizing that attention is a scarce resource, people select information guided by their utility. The idea of rational inattention stems from Information Theory laid out in the seminal work of Shannon (1948). Shannon (1948) provides a rigorous statistical definition of uncertainty about a random event X and then characterizes the maximum reduction of uncertainty about X that can be achieved through the knowledge of another event, Y. The uncertainty associated to an event X that takes value in $\{x_1, .., x_n\}$ with probability mass function p_X is given by the entropy of X:

$$H(X) = -\sum_{i=1}^{n} p_X(x_i)\log_2 p_X(x_i). \tag{9}$$

Expressions like (9)[7] provide a general and careful description of uncertainty associated to a random event based solely on its distribution.[8] To see how the initial uncertainty about X -$H(X)$- can be reduced, consider another random event, Y, with distribution p_Y and values in $\{y_1, .., y_m\}$. Let p_{XY} be the joint distribution of X and Y. Then, the maximum amount of uncertainty about X that can be reduced through the stochastic

[7] In (9), base 2 in the logarithm corresponds to calculating entropy in bits.

[8] Entropy satisfies four criteria that make it a universal measure of uncertainty: continuity, symmetry, maximality and additivity. The latter is probably the most appealing from the perspective of economic theory in that it relates the uncertainty of a system with the uncertainty of its subparts. This property makes entropy a suitable tool to evaluate uncertainty in complex economic systems.

knowledge of Y is given by their mutual information

$$I\left(p\left(\cdot_x,\cdot_y\right)\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}p_{XY}\left(x_i,y_j\right)\log_2\left(\frac{p_{XY}\left(x_i,y_j\right)}{p_X\left(x_i\right)p_Y\left(y_j\right)}\right)\tag{10}$$

where I have used $I\left(p\left(\cdot_x,\cdot_y\right)\right)\equiv I\left(X;Y\right)$ to make explicit that mutual information depends on the copula of the random event X and the random event Y. Mutual information measures dependence of two random events: it is positive and it is zero if and only if X and Y are independent. When two random variables are perfectly correlated, $I\left(p\left(\cdot_x,\cdot_y\right)\right)\rightarrow\infty$. Shannon (1948) prevents mutual information to go to infinity by stating that in any given period, the maximum amount of information about an event X that can be reliably retrieved through the knowledge of another random event Y using a communication channel is bounded by a number, that is, the channel capacity. This means that $I\left(p\left(\cdot_x,\cdot_y\right)\right)\leq\kappa$, where κ is the maximum number of bits of information that the channel can transmit. This seemingly abstract concept has a familiar correspondence in day-by-day experience: The maximum amount of information that we can download from our computer cannot exceed a number - the transmission rate- provided by the manufacturer. Likewise, we cannot instantaneously respond to a given E-mail. The amount of time it takes to respond to an E-mail depends on the content of the E-mail -i.e., initial uncertainty about the random variable X-, but also on how much information we want to process about the E-mail in order to produce a sensible response, Y. In the previous analogy the channel through which we transmit information from the original E-mail to the reply is our brain. The joint distribution $p\left(\cdot_x,\cdot_y\right)$ in (10) stands for the information we need to process about the content of the E-mail having in mind a reply, to make sure that the mapping between E-mail and reply is as accurate as we want. Finally, the upper bound on information transmission relates to the capacity of our brain to process information and to produce an action.

The latter analogy bears the salient elements of rational inattention and its use in economics. Consider a person who decides how much to consume and work while facing uncertainty about wages. Had he had infinite processing capacity, he would instantaneously map information about wages and wealth into an optimal policy for consumption and labor as in (6). Under information processing constraint à la Shannon, policies such as (6)-(8) are no longer feasible, for they require processing information at infinite rate. Introducing information processing constraints to the problem implies recognizing that it takes time and effort to map information about wealth accumulation -that depends on both savings and labor income, as equation (3) shows- into operative consumption and labor plans. As replying to an E-mail requires a person to turn attention to its content, *a fortiori* deciding labor and consumption with information processing constraints requires turning attention to current and future possibilities of savings and current and future possibilities of labor income. Moreover, in both the examples of E-mail and consumption/labor decisions, people choose the information about their state -i.e., E-mail received in the first case, wealth in the second case- having in mind how accurate they want their decision to be -i.e, reply to the E-mail and consumption/labor plans-. Finally, the accuracy of the reply of the E-mail and consumption/labor plans depends on how much effort we are willing to make in processing information. In turn, this effort hinges on the capacity of the channel -i.e., the brain- as much as the result of trading off of the cost of processing information for the utility that we derive by writing an adequate response and having a well specified saving and work policy.

2.2.1 Set up and Timing

I assume that the households in the economy share the same characteristics in terms of preferences, endowment and their capacity of processing information. This allows me to focus on a representative agent's economy. By assumption, although all the information is available, it cannot be fully processed by the consumer. This feature of the theory translates into the assumption that wealth is unknown. The agent enters the world with a belief about it, $b(W)$. He chooses a signal that conveys information jointly on his wealth (W) and decision ($A \equiv \{C, L\}$) of consumption, C, and work L. Let $p^*(\cdot_w; \cdot_a)$ denote the joint probability of wealth and decisions implied by the optimal choice of the signal. Note that the signal can provide information about any dimension of behavior - A- and wealth - W- that the person wants, with the restriction that the informativeness of the signal cannot exceed his processing capacity. Such a signal provides him with a rule of conduct for consumption and labor choices. In period $\tau = 1$, the household draw from the optimal choice $p^*(\cdot_w; \cdot_a)$ his consumption profile (c^*), and labor supply (l^*).

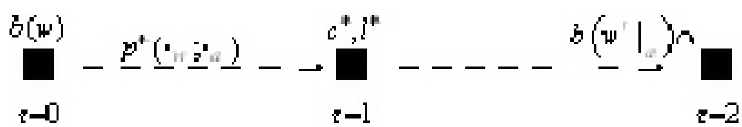

Figure 0: A typical day of a rationally inattentive person.

Then, he observes the outcomes of his choices and use the observation to update rationally his knowledge of wealth ($b(w'|_a)$). This complete his day. The day after he follows the same routine starting with $b(w'|_a)$ as his new prior. Figure 0 describes the events.

2.3 Statement and Recursive Formulation of Consumers' Problem

I discuss each element of the model in turn, starting from the constraints. First I present the budget constraint and discuss its role in updating the knowledge of wealth for an information constrained consumer. Next I turn to the information-flow constraint, key of the model. Finally I present the objective function and cast the problem into a recursive formulation.

The structure of the economy follows closely the one of Tutino (2008) to which I refer for the mathematical details. For completeness, Appendix A proves rigorously that the problem admits a recursive formulation and that the resulting Bellman equation is a contraction.

2.3.1 Budget Constraint and Update

Consumers maximize their lifetime utility function, defined over a consumption good and leisure. Let C denote the consumption good and L be labor. I collect the actions at time t in the set $\mathbb{A}_t \equiv \{C_s, L_s\}_{t \leq s \leq \infty}$.

Consumers are limited in their choices by the budget constraint

$$W_{t+1} = R_t \left(W_t - C_t\right) + s_t L_t \tag{11}$$

where $R_t = R$ is the (constant) interest rate on savings, $(W_t - C_t)$, s_t is the wage the agent receives in exchange of L_t units of labor. As in the setting with infinite processing capacity, the process characterizing the wages is a stationary i.i.d. Markov process. Note that the only source of stochasticity in the model comes from the wage rate, s. Moreover, people wish to reduce their uncertainty about the linear combination of savings and labor income as displayed in (11). Since information is fully and freely available, people can directly acquire signals about the law of motion of wealth although they cannot observe the exact value of wealth due to information-processing constraints. I assume that households make their consumption and labor choices knowing that the mean of the wages is fixed at $\bar{s}$. This knowledge is embedded in a prior, $g\left(w_t\right)$, over the possible realizations of wealth. As consumers go through life, they update rationally this belief with signals on wealth and the observation of their past behaviors.

Let $a_t \equiv \{c_t, l_t\}$ be a particular behavior of consumer at time t where c_t is a specific outcome of the random variable C and similarly l_t is a specific outcome of the random variable L. The choice a is drawn from the optimal choice $p^*\left(\cdot_w; \cdot_a\right)$.

The posterior of wealth conditional on observing a particular $a = \tilde{a}$ follows by Bayes' law :

$$g'\left(w'|_{\tilde{a}}\right) = \int_w T\left(w'; w, \tilde{a}\right) \cdot p\left(w|\, \tilde{a}\right) dw \tag{12}$$

where $T\left(w'; ., .\right)$ is the transition function commanded by the dynamic of wealth (11) [9] and $p\left(w|\, \tilde{a}\right)$ takes into account the potential noise in the current observation of the state. This noisy observation is carried over one period ahead to infer next period's state.

[9] Actually, the operator $T\left(w'; w, \tilde{a}\right)$ assigns probabilities to w' conditional on the value of $\tilde{a}$ and w. Since current values of w are not observable, the operator $T\left(.\right)$ applies an expectation over the unknown w. For a particular realization of $\tilde{a} = \left\{C = \tilde{c}, L = \tilde{l}\right\}$, the operators is defined as:

$$T\left(w'; w, \tilde{a}\right) \equiv E\left(W'\right) = R \int_w \left(w - \tilde{c}\right) dw + \bar{s}\tilde{l}$$

2.3.2 Choice Variable

Before processing any information about wealth (W), consumption (C) and labor (L) are random variables from the perspective of the consumers. To see why, first consider that the household cannot choose C and L optimally without relating wealth to behavior (C and L). Coming into the world with a probabilistic knowledge of W, mapping W into C and L translates into finding a joint relation among wealth, consumption and labor, that makes the information about wealth as related as possible to consumption and labor. The selection of information about wealth useful to consumption and labor, that is, the joint probability distribution of wealth, consumption and labor, $p(\cdot_w, \cdot_c, \cdot_l)$, is key in the optimization of the consumers since it affects current beliefs and their updates.[10]

To clarify this point, suppose that information flows at infinite rate. In this case, the optimal $p(\cdot_w, \cdot_c, \cdot_l)$ will be degenerate assigning to each value of w one value for $c^*(w)$ and $l^*(w)$, as in the solution (6)-(8). By contrast, suppose that processing information about wealth is too difficult for the consumer, then the consumer will be better off processing very limited amount of information about wealth. This is equivalent to choosing the minimal amount of information about wealth that allow the consumer to set c and l constant for each value of w without breaking his budget. Such a behavior implies that consumption and labor will be almost independent on wealth. When the information-processing effort lies in between this two extreme cases, optimizing consumers aim at setting $p(\cdot_w, \cdot_c, \cdot_l)$ such that the conditional probability of wealth given consumption and labor is as close to wealth as possible given the information constraint and the preference of the consumers.[11]

Once consumers choose an optimal policy, $p^*(\cdot_w, \cdot_c, \cdot_l)$, consumption and labor are drawn from that distribution.

2.3.3 Information Constraint

As explained before, people with limited processing capacity, select optimally information about wealth and behavior within their cognitive possibilities. I model the restriction that these cognitive possibilities are finite using Shannon's mutual information.[12] between the random variables W and A. This technology measures the maximum reduction in uncertainty associated to a system as difference between the initial uncertainty -entropy of W- and the knowledge of the variable W provided by the observation of A -i.e., conditional

[10] Alternatively, one can think of the choice of $p(\cdot_w; \cdot_c, \cdot_l)$ as equivalent to choosing a signal about wealth and behavior. The consumer decides the scope of the signal according to his preferences but the overall informativeness of the signal is constrained by his information-processing limits.

[11] Exploring the interaction between information processing constraints and general specifications for preferences is relatively novel to the literature of rational inattention which has focussed mainly on the Linear Quadratic Gaussian (LQG) framework (Sims 2003, Mackowiak and Wiederholt (2008a, 2008b). See Tutino (2009) for a discussion of the advantages of moving into a fully endogenous choice of signals with respect to the LQG framework.

[12] See Shannon (1954), Sims (2003, 2006).

entropy of W given A-. Since mutual information depends only on the joint distribution of W and A for a given belief, this way of modelling residual uncertainty is applicable without additional restrictions on the nature and characteristics of the channel. In my setting, Shannon capacity captures the ability of consumers to interpret news about their wealth, thereby regulating the speed of reaction of their behavior to these news.

I model people's ability to map information about wealth into consumption and labor decisions by assuming a constant and exogenous shadow cost on the information-processing constraint -i.e., mutual information between W and A-.[13] In the model, such a cost is denoted by θ. This assumption has the interpretation that mapping each bit of information about wealth into consumption and labor decisions costs the same processing effort to the consumers. Different from Sims (2003) and Mackowiak and Wiederholt (2008a) where the capacity is fixed and exhausted every period, fixing the shadow cost of processing information has the appealing property that consumers can effectively choose the amount of uncertainty they want to reduce each period according to their (perceived) financial conditions and their preferences. For instance, a person who finds it extremely costly to process information about wealth -i.e., high θ-, might choose to pay attention to wealth only after he observes that his consumption/leisure profile has changed significantly over time. By contrast, a person with relative better abilities to process information -i.e., low θ-, might find it optimal to keep close track of his wealth in order to enjoy combinations of consumption and leisure that maximize his utility.

Formally, let $\mathcal{I}\left(p\left(\cdot_w, \cdot_a\right)\right)$ be the mutual information implied by the choice of the joint distribution of W and A, $\left(p\left(\cdot_w, \cdot_a\right)\right)$. The constraint that limits the amount of processable information at each time t is given by :

$$\kappa_t = \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right) = \int p\left(w_t, a_t\right) \log\left(\frac{p\left(w_t, a_t\right)}{\left(\int p\left(\tilde{w}_t, a_t\right) d\tilde{w}_t\right) g\left(w_t\right)}\right) dw_t da_t \qquad (13)$$

The expression in (13) says that the maximum uncertainty that the consumer can reduce about his wealth through observation of consumption and labor supply is at most κ bits per unit of time. Mapping formulae into the intuition from the previous section, had the consumer had infinite processing capacity, he would have been able to choose a signal which makes each of his actions informative about a particular value of wealth. This results in a policy function for consumption, labor and wealth that depends on the -now observable- value of wealth. On the other extreme, with no processing capacity, the best one can do is to assign all the probability to a particular value of A. This choice makes the variables W and A independent of each other, $\left(\mathcal{I}\left(p\left(\cdot_w; \cdot_a\right)\right) \to 0\right)$. Every day, such a person spends the same amount of cash in consumption and the same amount of time working, regardless of his financial possibilities. In the intermediate case, if the

[13] Note that having a shadow cost associated to information processing and $\kappa = \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right)$ is isomorphic to assuming a bound on the maximum amount of capacity while having the constraint holding with inequality, that is $\bar{\kappa} \geq \kappa = \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right)$. The latter approach is adopted by, e.g., Sims (2003) and Mackowiak and Wiederholt (2008a) while the first approach is adopted by e.g., Sims (2006) and Tutino (2009).

person can process a finite amount of information, he attends to information that make his saving and labor decisions as related as to wealth as his utility commands and his information-processing constraint allows.

2.3.4 Objective

Household's problem is to maximize the infinite horizon expected utility of consumption and leisure discounted at factor β. Let θ be the (fixed and exogenous) shadow cost of processing information κ in (13). Moreover, let utility be specified by:

$$u(c,l) = \log c - \frac{\alpha}{\eta+1} l^{\eta+1} \tag{14}$$

where η is the inverse of Frisch elasticity of labor supply, ε, and $\alpha \in [0,1]$ is a constant disutility associated to labor.

The control for agent's maximization is a signal, $p(w,a)$ that solves : [14]

$$V(g(w)) = \max_{p(w,a)\in\mathcal{D}} \sum_{w}\sum_{a} u(c,l)\, p(w,a) - \theta\kappa + \beta\left(\sum_{w}\sum_{a}\left(V\left(g'\left(w'|_{a}\right)\right)\right) p(w,a)\right) \tag{15}$$

subject to

$$\kappa = \mathcal{I}\left(p\left(\cdot_w, \cdot_a\right)\right) = \sum_{w}\sum_{a} p(w,a) \log\left(\frac{p(w,a)}{\left(\sum_{\tilde{w}} p(\tilde{w},a)\right) g(w)}\right) \tag{16}$$

and (12) and the requirement that $p(w,a) \in \mathcal{D}$ where $\mathcal{D}$ is the set of all distributions that satisfy

$$\sum_{a} p(\tilde{w},a) = g(\tilde{w}) \tag{17}$$

$$0 \le p\left(\cdot_w;\cdot_a\right) \le 1,\ \forall (w,a) \tag{18}$$

$$\sum_{w}\sum_{a} p(w,a) = 1. \tag{19}$$

In addition (3) and (5) are imposed.

Taking first order condition with respect to $p(w,a)$ results in [15]

[14] For a formal prove that the infinite problem of the household can be written as a Bellman equation see Appendix A.

[15] Note that the first order condition in (20) is valid for $\theta > 0$. If $\theta = 0$, then the probabilities $g(w)$ and $p\left(\cdot_w,\cdot_a\right)$ are degenerate. In this case, $\mathcal{I}\left(p\left(\cdot_w;\cdot_a\right)\right) = 0$, and using Fano's inequality (Thomas and Cover 1991),

$$\begin{aligned} c\left(\mathcal{I}\left(p\left(\cdot_w;\cdot_a\right)\right)\right) &= c(w) \\ l\left(\mathcal{I}\left(p\left(\cdot_w;\cdot_a\right)\right)\right) &= l(w) \end{aligned}$$

which makes the first order conditions for this case the full information solutions (6)-(8).

$\partial p(w,a)$:

$$0 = u(c,l) + \beta V\left(\bullet|_a\right) + \theta\left(\log\left(\frac{p(w,a)}{\sum_{w''} p\left(w''; \cdot_a\right)}\right) - \sum_w \sum_a p(w,a) \frac{\partial\left(\sum_{w''}\left(w''; \cdot_a\right)\right)}{\partial\left(p(w,a)\right)}\right)$$

$$+\beta\left(\sum_w \sum_a \left[\frac{\partial\left(V\left(g'\left(w'|_{a_t}\right)\right)\right)}{\partial g\left(w'|_{a_t}\right)} \frac{\partial g\left(w'|_a\right)}{\partial p(w,a)}\right] p(w,a)\right) \tag{20}$$

where

$$\frac{\partial g\left(w'|_{a_t}\right)}{\partial p(w,a)} = \frac{\partial\left(\sum_w T\left(w'|\, w,a\right) p\left(w|\, a\right)\right)}{\partial p(w,a)}$$

$$= \sum_w \frac{T\left(w'|\, w,a\right)}{\sum_{a''} p\left(\cdot_w; a''\right) \tilde{p}(a)} + \sum_w \frac{\partial T\left(w'|\, w,a\right)}{\partial p(w,a)} p\left(w|\, a\right).$$

Define the solution to the optimization problem of the consumers as the distribution $p^*\left(\cdot_w; \cdot_a\right)$. The realized outcomes $\{c_t, l_t\}$ are then drawn from the optimal joint $p^*\left(\cdot_w; \cdot_c, \cdot_l\right)$. Appendix B derives the solution of a static, low-dimensional version of problem (15)-(19) that admits a quasi analytical solution.

3 Solution Strategy

The computation methodology follows closely Tutino (2009) to which I refer for technical details. However, there is a computational difference on the construction of the simplex that I shall highlight. This paper uses a uniform random grid to generate the simplex. Such a method is more efficient in terms of computational time than a non-uniform random grid and it requires less point to span the simplex. To sketch the methodology, we start with n as the number of possible values that w can assume. Then, each point of the simplex, Δ, is an $n-$array each of whose rows contains m random values belonging to the interval $[0,1]$. The distribution of values is uniform in the sense that it has the conditional probability distribution of a uniform distribution over the whole m-cube, given that the sum per row is 1. The algorithm randomly determines the placement of random points in the $n-1$ dimensional simplexes.

To map the finer state space into Matlab possibilities, I interpolate the value function with the new values of (12) using a kernel regression of $V(\cdot)$ into $g'\left(w'|_a\right)$. I use an Epanechnikov kernel with smoothing parameter $h = 3$.[16] For the partition on wealth, I choose an evenly spaced grid with 20 points where w takes values in $[1, .., 10]$. Similarly, the partitions on c and l have 20 points each and take values in $[0.53, .., 5.3]$ and $[1, .., 6]$ respectively.[17] The value iteration converges in about 220 iterations.

[16] The reason why I choose $h = 3$ comes from experimental trials for $h \in [2,5]$ with increments of 0.5 (*i.e.*, $h = 2$, $h = 2.5, ..$). While the results do not change substatially as I vary h in this range, the computational time is lower when I set $h = 3$.

[17] This discretization makes the joint distribution per simplex point a $400 * 20$ matrix.

Table 1 reports the benchmark parameter values.

Benchmark Values

	Discretization
$\bar{s}$	1.2
α	1.1
γ	1
η	1
R	1.02
Discount Factor, β	0.98

Table 1

In the calibration in Table 1, I assume that $R\beta = 1$ and log utility of consumption ($\gamma \to 1$). These two parametrization jointly imply consumers' desire to smooth consumption throughout their lifetime. Also, the benchmark model has Frisch elasticity of substitution equal to 1. Such a parametrization of Frisch elasticity is commonly used in the literature[18] and roughly consistent with macro evidence on labor supply elasticity. The specification of the preferences in the benchmark case implies that utility is concave and increasing in consumption and concave and decreasing in labor.

4 Predictions of the Model

The goal of this section is to provide suggestive evidence on how the model performs when compared to U.S. data. The data that I use for consumption are non durable goods from the Bureau of Economic Analysis while data for average hours worked are available from the Bureau of Labor Statistics. I present two sets of results in order to relate the predictions of the model to cyclical as well as secular movements of labor, consumption, wealth and wages. With respect to cyclical movements, I focus on a set of business cycle facts regarding consumption, labor and their volatility, persistence and comovement. The business cycle data have quarterly frequency from 1964.I to 2007.II. With respect to secular movements, I analyze the long-run consequences of changes in wages for consumption and labor supply.

4.1 Business Cycle Facts through Shannon's lenses.

I construct figures and statistics by de-trending the data with the HP filter, using a value of $\lambda = 1600$. I then used the detrended data to compute mean, standard deviation and correlations. I compare the results of the model with the detrended series under the observation that processing information through a Shannon channel filters out high frequency component of the variable(s) of interest (see Guo et al. (2005), Verdù (1996,

[18] See Shimer (2009) and references therein.

1999) and Sims (1998, 2003). The choice of HP filter over other filters is simply to ease the comparison with the business cycle literature.

As for the simulated series, I define a model as a set $M = \{\theta, \gamma, \eta\}$ and I compare the business cycle facts to specifications of the model that assume $\theta = 2, 0.2, 0.02$, $\gamma = 1$ and $\eta = 1$. I choose three values of θ as a proxy for three types of individuals that face three different shadow costs of processing information ranging from low ($\theta = 0.02$) to medium ($\theta = 0.2$) to high ($\theta = 2$). The choice of these particular numerical values can be explained as follows. I verify empirically that given the discretization of core states and core decisions and the baseline model with log utility of consumption and concave disutility in labor, a value of θ between 1 and 3 leads to the same quantitative results in terms of choice of distribution. Thus, I pick the middle value in the set $\theta \in [1, 3)$ for the high shadow cost of information-processing. The optimal choice of the joint distribution $p(w, a)$ is similar for value of $\theta \in (0.1, 0.6]$. Again, I pick the middle value in the interval for the second choice of θ. Similar reasonings conduct to $\theta = 0.02$ as lower value of θ when θ takes up values in $(0.01, 0.05)$. Notice that for values of θ above 5, households acquire very little information about wealth and set consumption and labor basically constant. Also, values of θ below 0.05 deliver a solution very close to the full information case.

To get a quantitative assessment of what these shadow costs mean in terms of loss in utility, I compute the average difference in lifetime utility between the infinite capacity case and each of the θ-cases considered. That is, I set $u(c, l) = \log(c) - (\alpha/(\eta + 1)) * l^{\eta+1}$ and I compute $E\left(u(c, l) - u\left(c^{\mathcal{I}}, l^{\mathcal{I}}, \theta\right)\right)$ where $E(u(c, l))$ is average lifetime utility under infinite capacity case while $E\left(u\left(c^{\mathcal{I}}, l^{\mathcal{I}}, \theta\right)\right)$ is the utility under a particular value of θ when θ takes values in $(\{2\}, \{0.2\}, \{0.02\})$. With an average value of $E(u(c, l)) \simeq 1.6$ under full information, $\theta = 2$ is associated to a loss of about 22% in lifetime utility, $\theta = 0.2$ implies a loss of about 11% while $\theta = 0.02$ delivers a loss of about 6%. The values used for this computation can be found in Table $10a - 10c$ in appendix B. I assume that the economy is populated evenly by those three types.[19]

The computations are as follows. For each θ, mean, standard deviation and correlation of the simulated series are calculated after I take averages of $10,000$ Monte Carlo runs and simplex points. The statistics for each of this series are in Tables $10a - 10c$. In tables $10d - 10f$ the same methodology produces the results for the cases $M = \{0.2, 3, 1\}$, $M = \{0.2, 1, 4\}$ and $M = \{0.2, 1, .25\}$. The results for the simulated series in Tables 2-4 and Table 6 are calculated by computing an arithmetic average of the series just described for $\theta = 2, 0.2, 0.02$. I do not filter out low frequency variations in the solution paths generated by the model since there are no exogenous shocks at high frequency. I refer to the business cycle facts occurring between 1964.I and 2007.II as BC and to the rational inattention predictions with consumption and labor as RI .

[19] As it is possible to see from the statistics for each θ in Table 10a-10f in Appendix B, the main results do not change significantly if instead of having an average of the results across θ, I assume an average value of θ.

4.1.1 BC Fact 1: The ratio of standard deviations of consumption and hours is 1.31

The model predicts excess volatility of consumption with respect to labor supply. This result depends on the bound of information-processing capacity and its interplay with risk aversion and disutility of labor. Consider a consumer with log utility in consumption and Frisch labor supply elasticity equal to 1. Had the household had infinite capacity, he would have chosen to smooth consumption by varying labor supply. With positive and finite information-processing capacity, he does not track wealth perfectly at high frequency. He selects a signal about wealth and changes labor and consumption according to the information content of the signal. Willing to smooth consumption, the household rationally chooses to save and work hard while he accumulates information about his financial possibilities. Once he realizes that he is rich, he increases consumption. However, to maintain a relatively high consumption profile for prolonged time, he keeps working hard. If he has accumulated too much savings -due to low informativeness of period by period signals- when the variation in consumption occurs is sizeable. Furthermore, such variation would be bigger than the one of labor to avoid taking risks on future wealth.

Table 2 compares the model $M(\theta,1,1) \equiv \frac{1}{3}\left(M(2,1,1)+M(0.2,1,1)+M(0.02,1,1)\right)$ to U.S. data. The mean of the wages is constant throughout the simulation at 1.2.

US Business Cycle : Quarterly Data (1964.I-2007.II), HP filter, , %

	Std.Dev
Non Dur. Consumption(C)	0.84
Hours (L)	0.64

Simulated Data, average across models

	Std.Dev
Consumption (c)	0.88
Labor (l)	0.66

Table 2: Statistical properties of US Business Cycle and Model

The first finding of the model, parallel to BC 1, is:

RC Finding 1. *For the model $M(\theta,1,1)$, the standard deviation of consumption over labor is 1.33. The volatility is higher the higher the shadow cost of information is. Moreover, the higher the shadow cost of information, the more sluggish consumption and labor are with sudden adjustment following accumulation of wealth. This joint behavior of consumption and labor is responsible for the high volatility of the two series.*

4.1.2 BC Fact 2: Non durable consumption and hours have 1^{st} order serial correlation higher than 80%

In the model, consumption and labor are more persistent the lower the information flow. In the latter case, it occurs also that contemporaneous consumption and labor lag wealth.

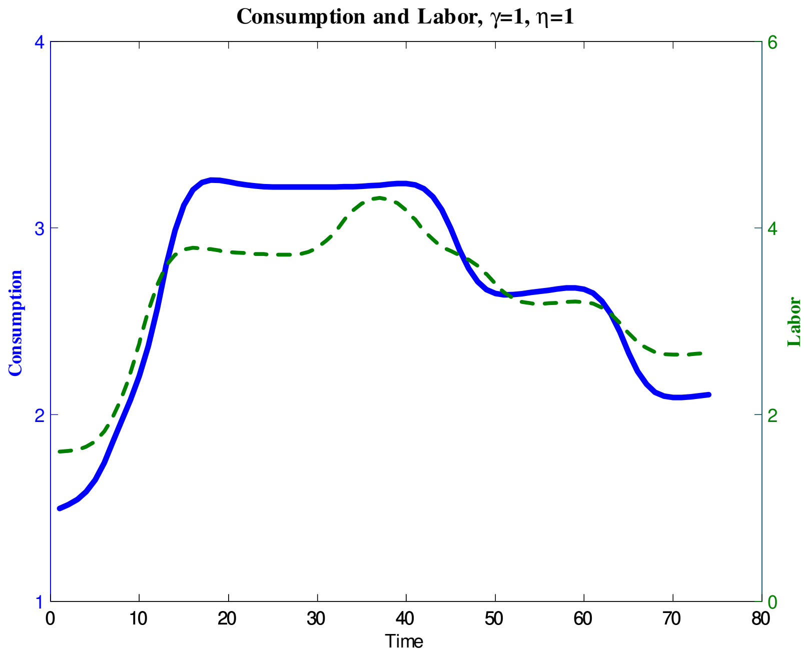

Figure 1: Sample Path of consumption and labor. Average across 10,000 Monte Carlo runs, M={θ,1,1}.

The intuition for these findings lies on the mechanism through which consumers update their knowledge of wealth, expressed in (12). Each period they choose a signal on wealth, decide consumption and labor based on the information from the signal and, given their choices, update their knowledge of wealth. The higher the processing cost, the less informative the signal. This in turn means that most of the update derives from the observations of past values of consumption and labor. As wealth accumulates, the signal consistently reports high values of wealth. This information triggers a reaction in behavior. The process leads to both delayed response to fluctuations of wealth and strong autocorrelation between current and past values of consumption and labor.

The comparison between model's findings and data is in Table 3.

US Business Cycle : Quarterly Data (1964.I-2007.II), HP filter, %

	Autocorr
Non Durable Goods(C)	93
Average Hours (L)	88

Simulated Data, average across models

	Autocorr
Consumption (c)	92
Labor (l)	87

Table 3: Statistical properties of US Business Cycle and Model

The counterpart for BC Fact 2 in the model can be summarized as:

RC Finding 2. *For the model $M(\theta, 1, 1)$, the autocorrelation of consumption and labor are above 80%.*

4.1.3 BC Fact 3: The contemporaneous correlation of consumption and hours is 78%.

US Business Cycle : Cross-Correlation, (1964.I-2007.II),HP, %

	$C(-1)$	C	$L(-1)$	L
C	0.93	1	0.77	0.78
L	0.66	0.78	0.88	1

Table 4:BC Fact 3

The model finds a strong comovement of labor and consumption. It also predicts a strong correlation between contemporaneous consumption and wealth when information flow is high. The reason for a strong relationship among contemporaneous values of wealth, labor and consumption in the model is twofold. First, the variables are related through the budget constraint (11) which is used by consumers to update their prior on wealth. Second, the optimal policy of the consumer delivers a tight link among these variables since it is the stationary joint distribution between choices (consumption and labor) and state (wealth).

Simulated Data: Cross-Correlation

	$c(-1)$	c	$l(-1)$	l
c	0.88	1	0.73	0.81
l	0.61	0.81	0.84	1

Table 5: Cross-correlation of consumption and labor. Average across 10,000 Monte Carlo runs, $M=\{\theta,1,1\}$

Table 5 and 6 show the comovement of labor and consumption and wealth, labor and consumption, respectively. While the fact that these variables are strongly correlated is robust, the size and sign of the correlation depend on the interaction between utility and information cost. As Table 10d in Appendix B shows, for a given shadow cost and disutility of labor, the higher the coefficient of risk aversion the higher the correlation between contemporaneous consumption and wealth. This result occurs because people that are relative more risk averse gather more information about wealth before changing their behavior than less risk averse people do. The reason is that people with high risk aversion want to avoid consuming too little when wealth is high.[20] Thus, they react to negative changes in wealth by varying labor supply and increase consumption in response to positive changes. If one is to fix the shadow cost and CRRA-coefficient, γ, and to increase Frisch elasticity of labor supply -from $\varepsilon = 1$ to $\varepsilon = 4$ or $\eta = 0.25$ in Table 10e in Appendix B-, then the contemporaneous correlation of labor and wealth increases while correlation between consumption and labor as well as the contemporaneous correlation between consumption and wealth decreases. The rationale for this result lies in the

[20] See Tutino (2009).

interaction between preferences and information constraint. People with high Frisch elasticity of labor supply acquire more information than people with relatively low Frisch elasticity -see Table $10f$ -. They do so because they want their labor supply to track closely wealth in order to prevent fluctuations of wealth to affect consumption. Thus, people with high Frisch elasticity trade off utils of consumption and leisure for being well informed about wealth fluctuations and, in turn, they are able to achieve a higher utility than people with low elasticity since they can maintain consumption constant throughout their lifetime -a desirable feature given the log utility of consumption-.

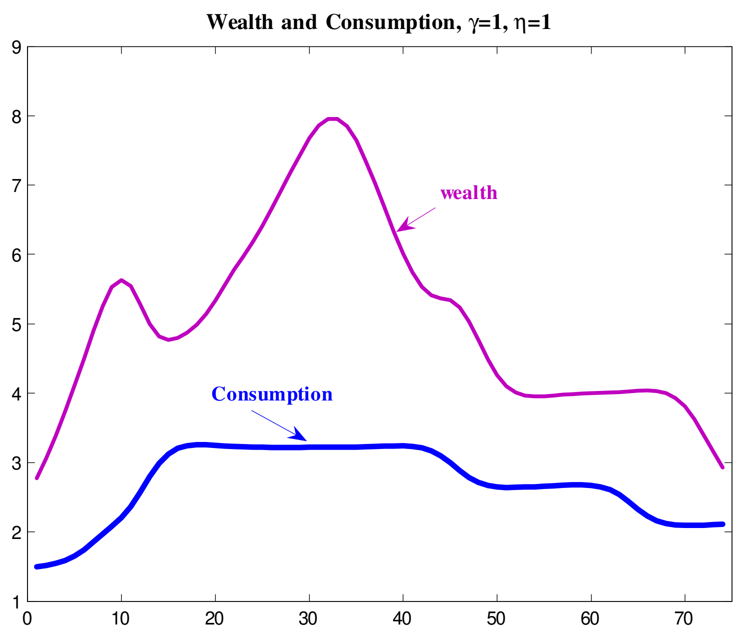

Figure 2: Sample Path of consumption and Wealth. Average across 10,000 Monte Carlo runs, M={θ,1,1}

The reason why lagged value of wealth are highly correlated with contemporaneous consumption comes from the interaction between the curvature of the utility function and the shadow cost of information. Concave utility together with high information cost trigger a conservative consumption profile and a consistent increase in consumption when the signal conveys information about high value of wealth (cfr. RC Finding 1).

With low elasticity of labor supply, the strong and positive comovement of consumption and labor (RC Finding 2) makes labor react in a way similar to consumption. When the elasticity of labor supply increases, people try to balance consumption smoothing and increase in leisure with the information available on wealth. The result is a weaker positive correlation of behavior and current and lagged values of wealth .

Cross-Correlation for Wealth

	$w\,(-1)$	w
c	0.73	0.61
l	0.68	0.58
w	0.73	1

Table 6: Cross correlation of wealth; average across $M\{\theta,\ \gamma{=}1{,}\eta{=}1\}$

The third fact from the model is:

RC Finding 3. *For the model $M\,(\theta,1,1)$ the contemporaneous cross correlation between consumption and labor is 0.81%. Moreover, consumption and labor are more correlated to lagged values of wealth than contemporaneous values of wealth.*

4.2 Growth Facts

Figures 3a-3b plot the response of labor and consumption to a 10% change in wage for the model (solid lines) as well as the infinite information equivalent (black dashed line). Both changes are assumed to occur in period $t = 1$ and they are known at $t = 0$. Figure 3a displays the responses of consumption and labor to a permanent change in wage, while Figure 3b show the responses of consumption and labor to a temporary change in wealth of the same amount.

The top panels of figure 3a present the impulse response for aggregate consumption (top right panel) and labor (top left panel). The bottom panels of figure 3a show responses of consumption (bottom right panel) and labor (bottom left panel) for individuals with 3 values of shadow costs of information constraint, i.e., $\theta = 2$ (green solid-star line), $\theta = 0.2$ (blue solid line) and $\theta = 0.02$ (magenta dotted-dashed line), together with the corresponding series with infinite information (black dashed line).

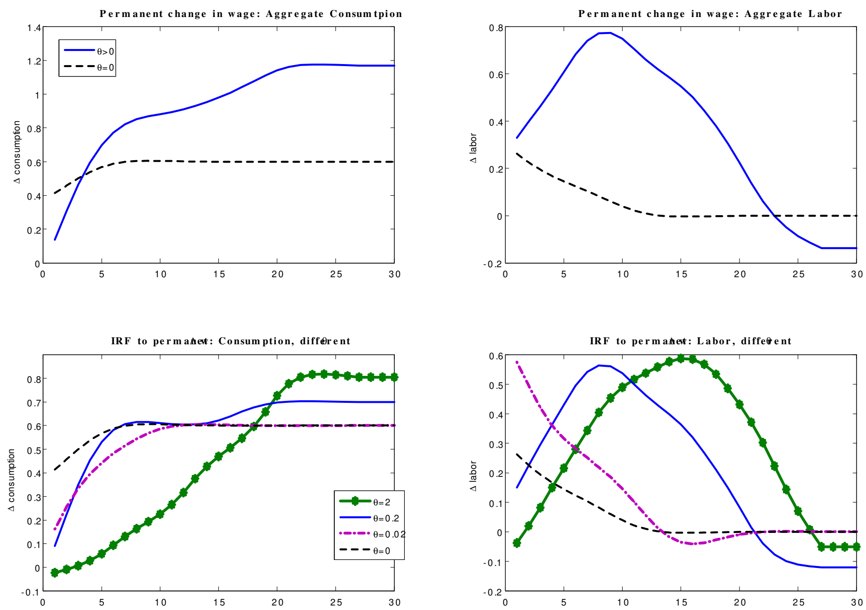

Figure 3a. Impulse response function for a permanent 10%−change in wage: consumption (first column) and labor (second column)

Consider the impulse response of consumption and labor for different values of θ (bottom panel). Under infinite information capacity, i.e., $\theta = 0$, consumption jumps quickly to its new steady state value while labor initially increases to sustain a higher level of consumption (wealth effect). Once consumption reaches its new higher value, the substitution effect commands a decrease in hours worked up until the household goes back to the work effort he had before the shocks. With shadow cost of processing information $\theta = 0.2$, people acknowledge the change in wealth slowly and cautiously increase both consumption and labor. As they wait to fully adjust their behavior to the increase in wages, their savings accumulate. Signals that wealth creeps up get stronger by the increase in savings and the increase in labor. Hence, type-$\theta = 0.2$ react to the those signal by moving consumption permanently up and slowly decreasing work effort. Note that the initial increase in work effort more than compensates for a permanently higher consumption. As a result, the steady state value of hours worked is lower than it was without the change in wealth. This mechanism is amplified for $\theta = 2$. In such a case, due to a low flow of information, these types are reluctant to change their behavior in response to the change in wealth. Such a reluctance results in more savings and, ultimately higher steady state consumption and higher hours worked with respect to the case with $\theta = 0.2$. Aggregating these types (top panels), when wages changes permanently, the model obtains a long run increase while the effect on labor is muted. This finding is consistent with secular patterns in U.S. data: in the long run, wages and consumption grow steady at about the same rate while movements in per capita hours are negligible.

Consider Figure 3b. It shows a temporary 10% change in wage. The first column of figure 3b shows impulse response functions for aggregate consumption (top panel) and individual consumption (bottom panel) when $\theta = 2$ (green solid-star line), $\theta = 0.2$ (blue solid line) and $\theta = 0.02$. The second column of figure 3b plots the corresponding impulse response for labor. In all the four pictures the black dashed line indicates impulse responses under the infinite information solution.

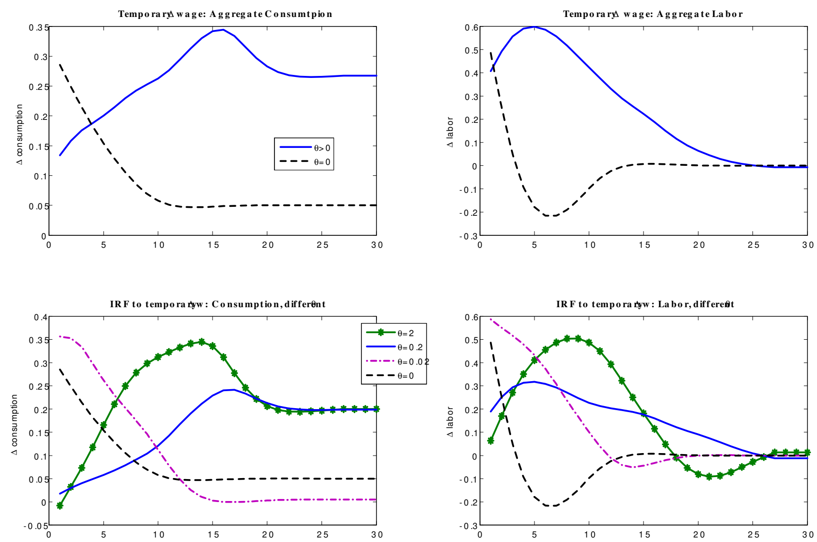

Figure 3b. Impulse response function for a temporary 10%−change in wage: consumption (first column) and labor (second column)

The impulse response for consumption and labor in the infinite information solution (black dashed line) show an initial increase in consumption and labor followed by a sudden decrease in hours worked while consumption reaches its new steady state. The model for $\theta = 0.2$ (blue solid line, bottom panels) predicts that consumption grows slowly following the shock to wage and so does labor supply. Accumulation of savings due to an increase in labor effort allows consumption to achieve an higher long run value. While consumption stabilizes to its new steady state level, substitution effect prevails and type $\theta = 0.2$ increases leisure, ending up to a new steady state value slightly lower than the one he enjoyed before the change in wage. People with $\theta = 0.02$ follows the same patterns as people with infinite processing constraint. However the effects of the temporary shock is much more persistent than in the full information case. People with $\theta = 2$ adjust their consumption and labor decision very slowly to the temporary shock. The logic of this result is akin to the one used for the permanent shock: income effect kicks in slowly due to low information flow and while people fail to react to the increase in wages they accumulate savings. Once people acknowledge the increase in savings, they adjust consumption so to keep it smooth from then on -recall that people have log utility-. At this point, the substitution effect prevails and people start enjoying more leisure. The aggregate impulse responses (top panels) confirm these patterns.

Studying the effect of a change in wage over the long run leads to the fourth finding of the paper:

RC Finding 4. *For the model $M(\theta, 1, 1)$ the long run effects of an increase in wages are: (1) a significant increase in consumption, (2) a negligible increase in hours worked. The effect on consumption are more pronounced the lower the information flow. The transitional dynamics of consumption and labor in the model have much more endogenous persistence compared to the infinite information case.*

5 Optimal conditional distribution

Figures 4a and 4b show the optimal conditional distribution of consumption and labor[21] respectively for a given value of wealth - $w = 1$, top panel $w = 4$, medium panel and $w = 8$, bottom panel- and two values of the shadow cost of information processing - $\theta = 0.2$, blue bars, and $\theta = 2$, solid green line-. The figures plot the optimal conditional distribution for a given simplex point.[22] Consider first the optimal conditional distribution of consumption when wealth is low -Figure 4a, top panel-. For $w = 1$, the optimal signal acquired by a person with relatively high information flow -$\theta = 0.2$, blue bars in the picture- places high probability mass on low values of consumption ($c = 0.7$) but he also allows for the possibility of higher consumption - $c = 2.3$ and $c = 3.1$- sustained by labor income. In fact, as the top panel of Figure 4b shows, the optimal distribution of labor conditional to $w = 1$ for a person with $\theta = 0.2$ places more than 60% probability on values of hours worked above the median level - $l = 3.5$ - of the support of labor. By contrast, when wealth is high -$w = 8$, bottom panel in Figure 4a and 4b- the agent with $\theta = 0.2$ assigns high probability to high values of consumption -$c = 5.1$-. Note that although the agent reduces his labor effort with respect to the case $w = 1$, he still places more than 40% of probability of working at and above the median value for hours so that he can maintain high value of consumption with labor income and savings. For a medium value of wealth, $w = 4$, the agent with $\theta = 0.2$ assigns most of the probability mass to values of consumption between 2 and 3. However, the agent assigns also positive -even if small- probability to high value of consumption -$c = 5.5$ with probability 0.02- counting on financing consumption expenditures through labor income. Also, for this type of agent, choices of labor are focussed on the medium values of the support with a

[21] The optimal conditional distributions plotted in Figures 4a and 4b are calculated as follows. For a given value of wealth, $w = w^*$, the optimal conditional distribution of consumption is:

$$p(c|\, w^*) = \sum_{l \in L} p^*(c, l|\, w^*)$$

where $p^*(c, l|\, w^*)$ comes from the solution of the optimization problem (15)-(19) and L is the set of possible values that labor, l, assumes. A similar expression holds for $p(l|\, w^*)$.

[22] In Appendix B, the values reported for the statistcs of consumption, labor, wealth and information flow are average across simplex-points.

peak at $l = 2.8$ whose probability is 0.42.

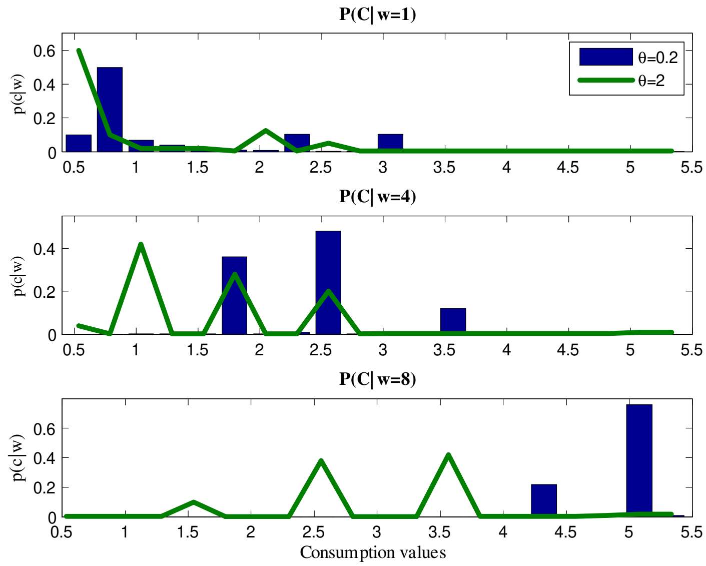

Figure 4a: Optimal conditional distribution of consumption for θ=0.2 (bar) and θ=2 (line)

Now consider an agent with the same preferences as the previous one but higher shadow costs of processing capacity, i.e., $\theta = 2$. By looking at figures 4a and 4b, it is evident that people with $\theta = 2$ have more spread distributions of consumption and labor than their $\theta = 0.2$ counterpart. The noisier behavior of their labor and consumption is due to the fact that people with $\theta = 2$ have a lower reduction in uncertainty about wealth than people with $\theta = 0.2$. Thus, their optimal probability is less informative than that of types $\theta = 0.2$. For instance, consider the conditional distribution of consumption and labor when wealth is high, $w = 8$. The person with $\theta = 2$ places higher probability on low values of consumption than the person with $\theta = 0.2$ does.

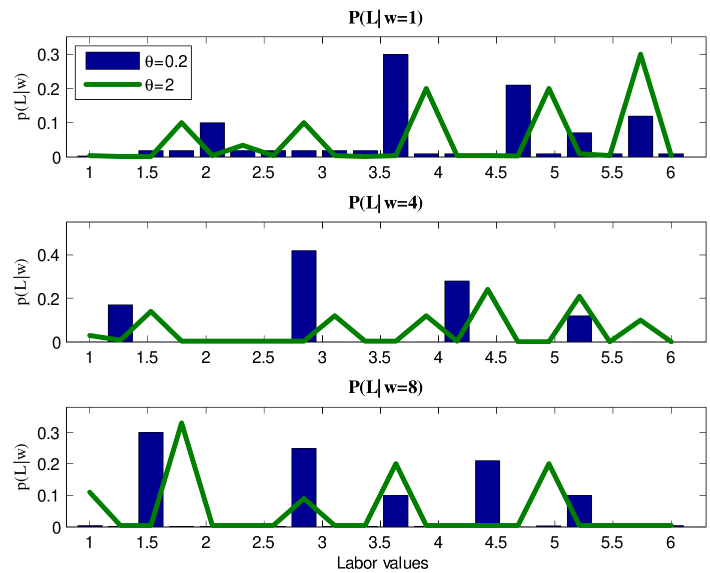

Figure 4b: Optimal conditional distribution of labor for θ=0.2 (bar) and θ=2 (line)

As tables 10a-10b in Appendix B reveal, the expected values[23] of consumption and labor for a person with $\theta = 0.2$ is 3.95 with variance 1.10 for consumption and 2.68 and 1.06 for labor. A person with $\theta = 2$ enjoys on average 3.55 units of consumption with variance 1.79 and works an average of 3.05 hours with variance 1.52. The average information flow is 1.08 bits for an individual with $\theta = 0.2$ and 0.73 bits for a person with the same preferences as the previous one but with $\theta = 2$.

5.1 Robustness

Tables $10d - 10f$ show the relationship between risk aversion, γ, Frisch elasticity of substitution, $1/\eta$, and shadow cost of information, θ. As table $10d$ shows, for a given θ and elasticity of substitution, the higher the coefficient of risk aversion , the higher the mean and the lower the variance of consumption. This finding makes intuitive sense since a risk averse household would save a lot during the early stages of life to enjoy high consumption later on due to the accumulated savings. Savings come from both low consumption and high labor supply at early stages of life triggered by the fear of running out of wealth. Once consumers realize they have built a sufficient buffer to cover for consumption and leisure expenses, they increase consumption and reduce -though by a lower extent- labor supply. The peak in consumption for these types of households

[23] All the statistics in Appendix B are evaluated as average across simplex points, e.g. :

$$E\left(C\right) = \sum_i \left(\sum_j \sum_k c_k * \left[p^* \left(c_k | \, w_j \right) g_i \left(w_j \right) \right] \right)$$

where $g_i\left(w\right)$ is a single simplex point.

occur later in their lifetime and labor supply is higher at the beginning than it is later on. Thus, information costs enhance precautionary savings. For a given θ and γ, a lower the Frisch elasticity of substitution (from $\eta = 0.25$ to $\eta = 4$ in tables $10e$ and $10f$, respectively) generates lower mean and lower variances for both consumption and labor. Keeping the degree of risk aversion fixed, a low elasticity of substitution for labor supply increases the income effect over the substitution effect. However, the presence of information processing constraint still favors substitution effect mitigating the income effect. If the signal on wealth is very noisy, consumers supply more labor than they would do in the case of perfect information and $\eta = 4$, since they are not certain that their wealth is actually decreasing. When the information collected signals that the wealth has been increased, labor supply suddenly decreases. The opposite occurs when consumers receive more and more information about a decrease in wealth.

To get a sense on how the shadow cost of information affects consumption and labor behaviors when the Frisch elasticity of substitution goes to infinity, consider $M(\theta) \equiv \{\theta; (\gamma \rightarrow 1, \eta = 0)\}$ where $\theta = 2, 0.2,$ 0.02. The first observation is that as the information costs increases, average consumption, labor and information flow decrease, while the standard deviations of these series increases. This is also true for wealth. These results are intuitive. Under full information, the characteristics of the utility function ($\gamma \rightarrow 1$) command a consumption profile smooth throughout their life-time. Moreover, with infinite Frisch elasticity of substitution, (i.e., $\eta = 0$) labor supply adjusts according to wealth fluctuations to accommodate consumption smoothing. When information flows at finite rate, rational households choose signals about wealth with the same purpose. If processing-information has low cost, $\theta = 0.02$, consumers can choose a signal about wealth informative enough to allow them to use labor supply to smooth fluctuations in wealth and, in turn, consumption. On the other hand, if information is costly, $\theta = 2$, consumers keep track of wealth slowly and, as a consequence, modify consumption and labor sporadically. When they do change their behavior, they do so by a significant amount. The resulting path for wealth inherits the higher variance of consumption and labor and, on average, has higher mean than in the previous case due to the increase in savings in periods of inertial behavior. A sample path of consumption under different $\theta' s$-scenarios is in Figure 5.

Comparing consumption for $u(c,l)=\log(c)-\alpha l$

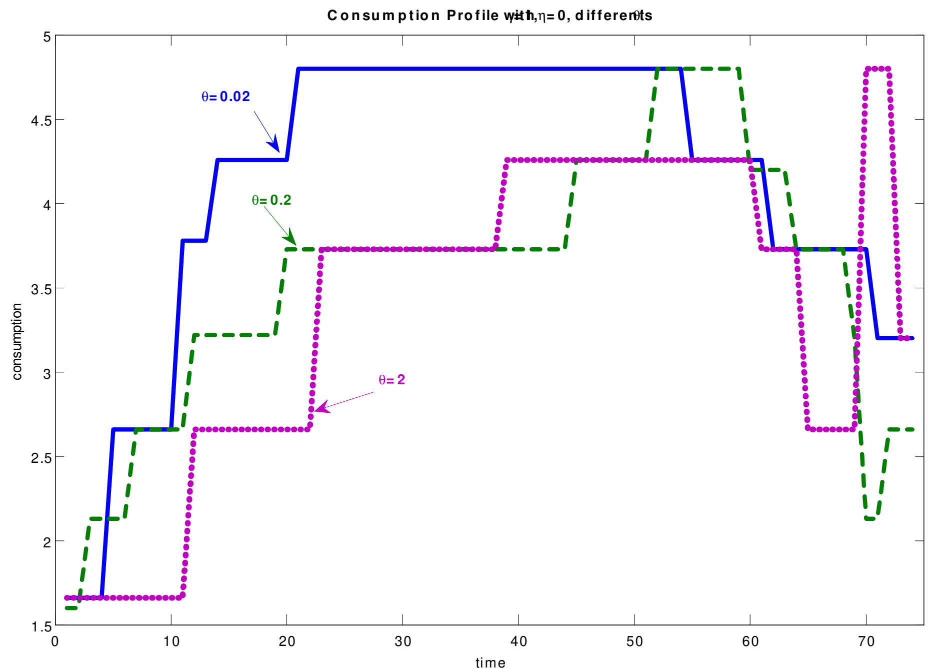

Figure 5: Sample path of consumption for different θ's.

Figure 6 confirms the intuition that consumption is smoother the lower information costs are. Consumers with $\theta = 0.02$ save at the beginning of the simulated period to enjoy high level of consumption later on. By contrast, consumers with $\theta = 2$ track with difficulties their wealth and this is reflected in a prolonged period of savings while processing information about wealth. This behavior results into slow and sizeable adjustments of consumption during the simulated period. One point worth attention is the existence of precautionary savings generated by information flow constraints. Individuals with less processing capacity ($\theta = 2$) push forward an increase in consumption more than the other people ($\theta = 0.2$ and $\theta = 0.02$) do. Types $\theta = 2$ acknowledge the accumulation of wealth due to the additional savings later in the simulation. This forces them to increase their consumption for a short period of time at the end of the simulation period.

Given the strong correlation between consumption and labor and the preferences of the individuals, it is not surprising that people with $\theta = 0.02$ work harder at the beginning of the simulation to finance their consumption, though they manage to enjoy some leisure at the end of the simulation (see Figure 7).

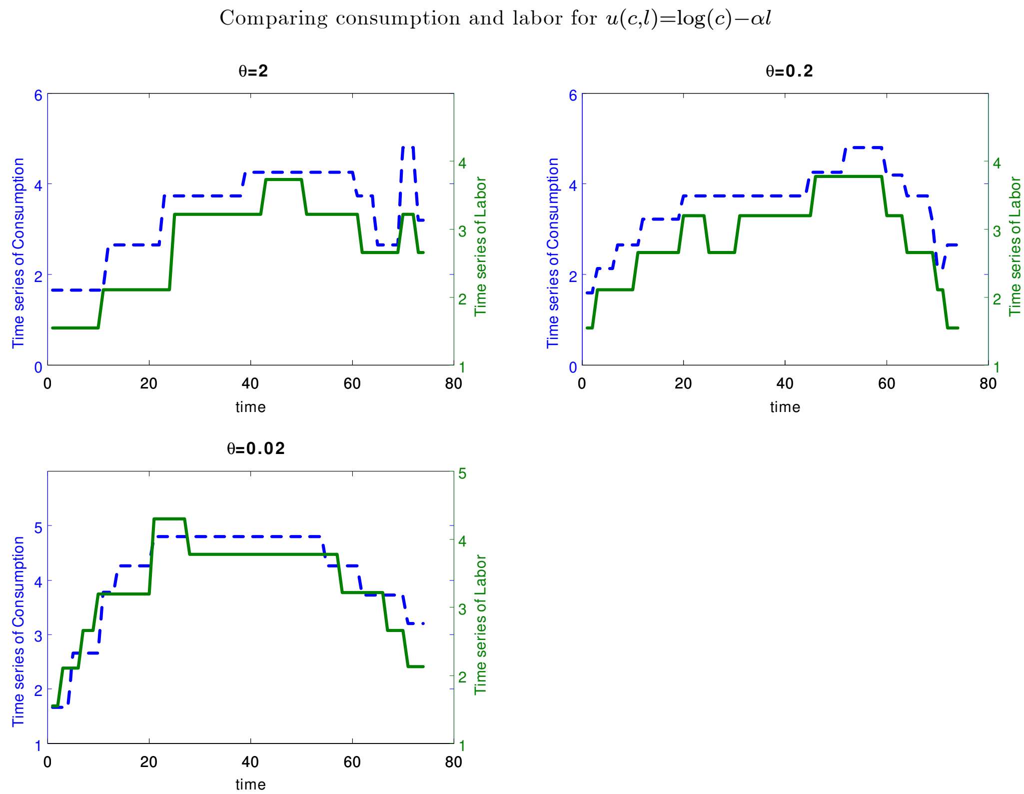

Figure 6: Time path consumption and labor, various θ

Correlation between consumption and labor is higher the higher the information costs. The intuition for this result is that the reaction of both consumption and labor behaviors to accumulation of wealth are delayed by individual's capacity of processing information. As they have better knowledge of how much wealth they have, they review both plans. People actions are mirrored in wealth accumulation. Individuals with $\theta = 0.02$ build up savings at the beginning of the period to dissave gradually later on. This is akin to consumption smoothing under full information.

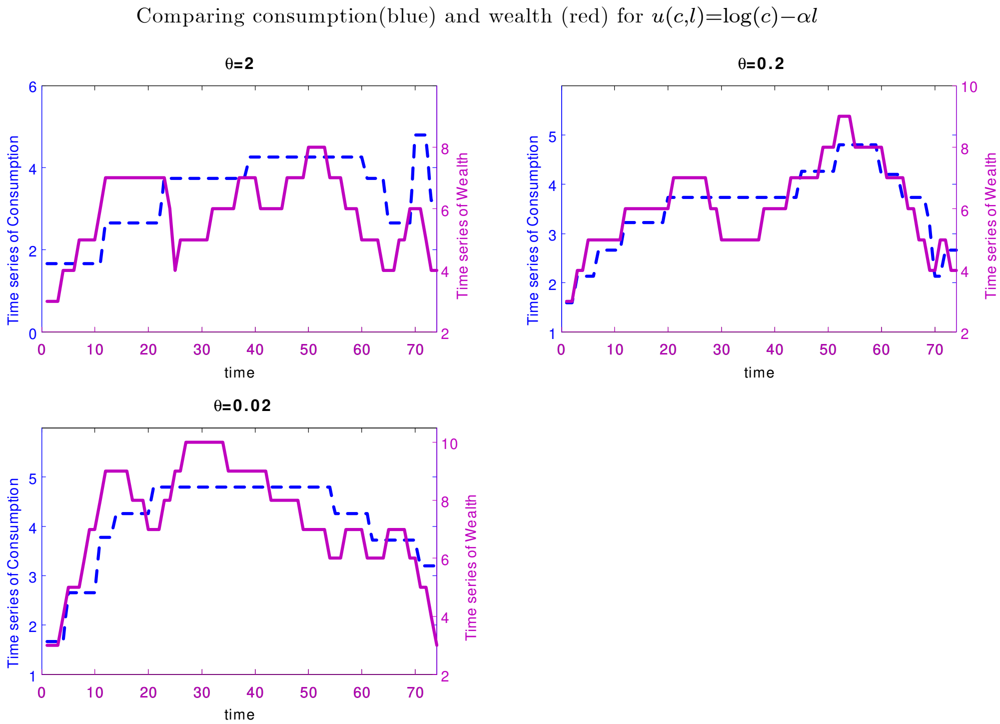

Figure 7: Time path wealth and consumption, log-lin utility, various θ

People who are more constrained in their choice of signals, adjust with delays consumption to fluctuations in wealth. Such delays smooth consumption while consumers are processing information but at the same time, call for major adjustments afterwards. Note also how consumption and labor lag wealth for $\theta = 2$. The cross-correlation coefficients between lagged wealth and current consumption is 0.65 while the contemporaneous correlation is only 0.47. A similar result holds for labor and lagged values of wealth. This finding is also consistent with intuition. Every period households receive little information about their wealth and rely on past values of consumption and labor to update their knowledge. While waiting, wealth accumulates and so does information until the consumers are convinced to change their behavior. This mechanism implies that behavioral response to movement in wealth is lagged. Finally, Figures 5-7 (together with Figures 8-11 in Appendix B) illustrate also the high persistence of the series documented in the previous section for $M(\theta) \equiv \{\theta; (\gamma \to 1, \eta = 1)\}$ whose statistics are in Tables $10a - 10c$. Not surprisingly, the persistence is higher the higher the information cost is.

6 Conclusions

I presented a model in which rational households optimize their lifetime utility under information-processing constraint à la Sims. I show that such a model, even in its simplicity, is able to replicate many empirical regularities of U.S. business cycle data as well as secular movements: higher volatility of consumption with respect to labor, persistence and strong comovement of consumption and employment, lagged response of consumption and labor with respect to wealth. Following a temporary as well as permanent increase in wages, the model produces a permanent increase in consumption and no significant changes in the long run growth of labor supply. Consistent with the empirical evidence, in the model the effects of a shock to wage spread out slowly through time. The main source behind the persistence of such a shock to wage is consumers' endogenous choice of information. When households face information-processing constraints, they select signals on their wealth and make consumption and labor decisions based on those signals. Each period, wealth evolves because of savings which depends on consumption and income. The latter in turn depends on labor supply and an exogenous stochastic wage whose distribution is fixed and known. Since movements in income affect the growth of wealth and consumers keep track of it by signals, the less informative the signals, the more persistent their choices. Once wealth accumulates and households acknowledge this growth through information collected, they change their behavior consistently. These predictions of the model are consistent with secular movements in labor and consumption in the U.S with long run per capita consumption exhibiting steady growth whereas per capita hours worked showing negligible growth. Moreover, the same mechanism is able to capture persistence and lags of the main macroeconomic variables over the U.S. business cycle. The findings of the paper suggest that making a leap to a fully rational inattention setting is worth the computational effort. For it gets us closer to understand and interpret empirical regularities in the U.S. data than the current theoretical macro literature.

Acknowledgement 1 *I am deeply indebted to Chris Sims whose countless suggestions, shaping influence and guidance were essential to improve the quality of this paper. I thank Ricardo Reis for valuable advice, enduring and enthusiastic support. I would also like to thank Nobu Kiyotaki, Philippe-Emanuel Petalas, Charles Roddie and Sam Schulhofer-Wohl. Any remaining errors are my own.*

References

[1] Aguiar, Mark, and Erik Hurst, 2005. "Consumption versus Expenditure." Journal of Political Economy. 113 (5): 919–948.

[2] Barro, Robert J., 1977. "Long-term contracting, sticky prices, and monetary policy." Journal of Monetary Economics. 3 (3): 305–316.

[3] Becker, Gary S., 1965. "A Theory of the Allocation of Time." The Economic Journal. 75 (299): 493–517.

[4] Blanchard, Olivier J., and Jordi Galı', 2006. "A New Keynesian Model with Unemployment." MIT Mimeo.

[5] Browning, M., Hansen L.P., and J. Heckman (1999), "Micro data and general equilibrium models", in: J.B. Taylor and M. Woodford (eds.), Handbook of Macroeconomics, vol. 1, Elsevier, 103-204.

[6] Chang, Y., and S. Kim (2005). "On the aggregate labor supply", Federal Reserve Bank of Richmond Economic Quarterly, 91(1), 21-37.

[7] Chang, Y., and S. Kim (2006). "From individual to aggregate labor supply: A quantitative Analysis on a heterogeneous agent macroeconomy", International Economic Review, 47, 1-27.

[8] Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, 2007. "Business Cycle Accounting." Econometrica. 75 (3): 781–836.

[9] Chen, Kaiji, Ayse Imrohoroglu, and Selahattin Imrohoroglu, 2007. "A Quantitative Assessment of the Decline in the U.S. Savings Rate." University of Southern California Mimeo.

[10] Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin, 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." Journal of Monetary Economics. 46 (2): 281–313.

[11] Gali', Jordi, Mark Gertler, and J. David Lopez-Salido, 2007. "Markups, Gaps, and theWelfare Costs of Business Fluctuations." The Review of Economics and Statistics. 89 (1): 44–59.

[12] Galı', Jordi, and Paul Rabanal, 2004. "Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar US Data?." NBER Macroeconomics Annual. 19: 225–318.

[13] Gertler, Mark, and Antonella Trigari, 2006. "Unemployment Fluctuations With Staggered Nash Wage Bargaining." NBER Working Paper 12498.

[14] Guo, D., S. Shamai , and S. Verdú, 2005 "Mutual information and minimum mean-square error in Gaussian channels," IEEE Trans. Inf. Theory, vol. 51, no. 4, pp. 1261–1283, Apr. 2005

[15] Hall, Robert E., 1997. “Macroeconomic Fluctuations and the Allocation of Time.” Journal of Labor Economics. 15 (1): 223–250.

[16] Hall, Robert E., 2005. “Employment Fluctuations with Equilibrium Wage Stickiness.” American Economic Review. 95 (1): 50–65.

[17] Hansen, Gary D., 1985. “Indivisible Labor and the Business Cycle.” Journal of Monetary Economics. 16 (3): 309–327.

[18] Heckman, J. (1993), "What has been learnt about labor supply in the past twenty years", American Economic Review, 83, 116-21.

[19] Imai, Susumu, and Michael P. Keane, 2004. “Intertemporal Labor Supply and Human Capital Accumulation.” International Economic Review. 45 (2): 601–641.

[20] Kaldor, Nicholas, 1957. “A Model of Economic Growth.” The Economic Journal. 67 (268): 591–624.

[21] Kalai, Gil. A subexponential randomized simplex algorithm (extended abstract).In Proceedings of the twenty-fourth annual ACM symposium on Theory of computing, pages 475–482. ACM Press, 1992.

[22] Kalai, Gil. Linear programming, the simplex algorithm and simple polytopes. Math. Programming, 79(1-3,Ser. B):217–233, 1997.

[23] Krusell, Per, and Anthony A. Smith, Jr., 1998. “Income and Wealth Heterogeneity in the Macroeconomy.” Journal of Political Economy. 106 (5): 867.

[24] Lovejoy, W. Computationally feasible bound for partially observable Markov Decision problem, Operations Research, 39:162:175,

[25] Lucas, Robert E. Jr., and Leonard A. Rapping, 1969. “Real Wages, Employment, and Inflation.” The Journal of Political Economy. 77 (5): 721–754.

[26] Mackowiak, B., and M. Wiederholt, 2008a, "Optimal Sticky Prices under Rational Inattention", American Economic Review, forthcoming

[27] Mackowiak, B., and M. Wiederholt, 2008b, "Business Cycle Dynamics under Rational Inattention", Mimeo.

[28] MaCurdy, Thomas E., 1981. “An Empirical Model of Labor Supply in a Life-Cycle Setting.” Journal of Political Economy. 89 (6): 1059–1085.

[29] Mankiw, N. Gregory, Rotemberg, J. Julio and Lawrence H. Summers, 1985, "Intertemporal substitution in Macroeconomics", The Quarterly Journal of Economics, MIT Press, vol. 100(1), pp. 225-51.

[30] Mankiw, N. Gregory, 1989. “Real Business Cycles: A New Keynesian Perspective.” Journal of Economic Perspectives. 3 (3): 79–90.

[31] McGrattan, Ellen R., and Edward C. Prescott, 2007. "Technical Appendix: Unmeasured Investment and the Puzzling U.S. Boom in the 1990s." Federal Reserve Bank ofMinneapolis Staff Report 395.

[32] Mertens, Karel, and Morten O. Ravn, 2008. "The Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks: Theory and Empirical Evidence." Cornell University Mimeo.

[33] Modigliani, Franco, 1977. "The Monetarist Controversy or, Should We Forsake Stabilization Policies?." American Economic Review. 67 (2): 1–19.

[34] Mortensen, Dale T., and Christopher A. Pissarides, 1994. "Job Creation and Job Destruction in the Theory of Unemployment." The Review of Economic Studies. 61 (3): 397–415.

[35] Mulligan, Casey B., 2002. "A Century of Labor-Leisure Distortions." NBER Working Paper 8774.

[36] Parkin, Michael, 1988. "A Method for Determining Whether Parameters in Aggregative Models Are Structural." Carnegie-Rochester Conference Series on Public Policy. 29: 215–252.

[37] Pissarides, Christopher A., 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." The American Economic Review. 75 (4): 676–690.

[38] Pissarides, Christopher A., 2007. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?." LSE Mimeo.

[39] Prescott, Edward C., 2004. "Why Do Americans Work So Much More Than Europeans?." Federal Reserve Bank of Minneapolis Quarterly Review. 28 (1): 2–13.

[40] Ragan, Kelly S., 2006. "Taxes, Transfers, and Time Use: Fiscal Policy in a Household Production Model." Ph.D. thesis, University of Chicago.

[41] Rogerson, Richard, 1988. "Indivisible Labor, Lotteries and Eequilibrium." Journal of Monetary Economics. 21 (1): 3–16.

[42] Rogerson, Richard, 2007. "Taxation and Market Work: Is Scandinavia an Outlier?." Economic Theory. 32 (1): 59–85.

[43] Rogerson, Richard, and Johanna Wallenius, 2007. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." NBER Working Paper 13017.

[44] Romer, Christina D., and David H. Romer, 2007. "A Narrative Analysis of Postwar Tax Changes." University of California, Berkeley Mimeo.

[45] Rotemberg, Julio J., and Michael Woodford, 1991. "Markups and the Business Cycle." NBER Macroeconomics Annual. pp. 63–129.

[46] Rotemberg, Julio J., and Michael Woodford, 1997. “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy.” NBER Macroeconomics Annual. 12: 297–346.

[47] Shannon, C. 1948. "A Mathematica Theory of Communication". Bell System Technical Journal, vol. 27, pp. 379–423 and 623–656, July and October, 1948

[48] Shimer, R. 2009. "Convergence in Macroeconomics: The Labor Wedge." American Economic Journal: Macroeconomics, 1(1): 280–97.

[49] Sims, Christopher A., 1998, "Stickiness", Carnegie-rochester Conference Series On Public Policy, 49(1), 317–356.

[50] Sims, Christopher A., 2003, "Implications of Rational Inattention", Journal of Monetary Economics, 50(3), 665–690.

[51] Sims, Christopher A., 2006 "Rational Inattention: A research Agenda", American Economic Review Paper and Proceedings.

[52] Tutino, Antonella, 2008 "The Rigidity of Choices, Lifetime savings under information processing constraints". Princeton Universtiy, Mimeo.

[53] Uhlig, Harald, 2003. “How Well do we Understand Business Cycles and Growth? Examining the Data with a Real Business Cycle Model.” inWolfgang Franz, Hans J¨urgen. Ramser, and Manfred Stadler (ed.), Empirische Wirtschaftsforschung: Methoden und Anwendungen, vol. 32 of Wirtschaftswissenschaftliches Seminar Ottobeuren, pp. 295–319.

[54] Verdú, S. 1996 "The Exponential Distribution in Information Theory," Invited Paper, Problemy Peredachi Informatsii, vol. 32, no.1, pp.100-111, Jan.-Mar. 1996 (In Russian). English version in Problems of Information Transmission, vol. 32, no. 1, pp. 86-95, Jan.-Mar. 1996.

[55] Verdú, S. 1999 , Information Theory: Fifty Years of Discovery, IEEE Press, Piscataway, NJ, S. Verdu and S. McLaughlin, Eds.

7 Appendix A: Bellman Recursion and its properties

This appendix follow closely the work of Tutino(2008). It establishes the main properties of the Bellman recursion in the discrete Rational Inattention consumption-labor model.

7.1 The Bellman Recursion is a Contraction Mapping.

Proposition 1. *For the discrete Rational Inattention consumption labor value recursion H and two given functions V and U, it holds that*

$$||HV - HU|| \leq \beta\, ||V - U||\,,$$

with $0 \leq \beta < 1$ and $||.||$ the supreme norm. That is, the value recursion H is a contraction mapping.

Proof. The H mapping displays:

$$HV\,(g) = \max_{p} H^{p}V\,(g)\,,$$

with

$$H^{p}V\,(g) = \left[\sum_{w\in W}\left(\sum_{a\in A} u\,(c,l)\,p\,(a|w)\right)g\,(w) - \theta\kappa + \beta\sum_{w\in W}\sum_{a\in A}\left(V\,(g'_{a}\,(\cdot))\right)p\,(a|w)\,g\,(w)\right].$$

Suppose that $||HV - HU||$ is the maximum at point g. Let p_1 denote the optimal control for HV under g and p_2 the optimal one for HU

$$HV\,(g) = H^{p_1}V\,(g)\,,$$
$$HU\,(g) = H^{p_2}U\,(g)\,.$$

Then it holds

$$||HV\,(g) - HU\,(g)|| = H^{p_1}V\,(g) - H^{p_2}U\,(g)\,.$$

Suppose *WLOG* that $HV\,(g) \leq HU\,(g)\,.$ Since p_1 maximizes HV at g , I get

$$H^{p_2}V\,(g) \leq H^{p_1}V\,(g)\,.$$

Hence,

$$\begin{aligned}
||HV - HU|| &= \\
||HV\,(g) - HU\,(g)|| &= \\
H^{p_1}V\,(g) - H^{p_2}U\,(g) &\leq \\
H^{p_2}V\,(g) - H^{p_2}U\,(g) &= \\
\beta\sum_{w\in W}\sum_{a\in A}\left[(V^{p_2}\,(g'_{a}\,(\cdot))) - (U^{p_2}\,(g'_{a}\,(\cdot)))\right]p_2 g\,(w) &\leq \\
\beta\sum_{w\in W}\sum_{a\in A}\left(||V - U||\right)p_2 g\,(w) &\leq \\
\beta\,||V - U||\,.
\end{aligned}$$

Recalling that $0 \leq \beta < 1$ completes the proof. ■

7.2 The Bellman Recursion is an Isotonic Mapping

Corollary *For the discrete Rational Inattention consumption-laving value recursion H and two given functions V and U, it holds that*

$$V \leq U \Longrightarrow HV \leq HU$$

that is the value recursion H is an isotonic mapping.

Proof. Let p_1 denote the optimal control for HV under g and p_2 the optimal one for HU

$$HV(g) = H^{p_1}V(g),$$
$$HU(g) = H^{p_2}U(g).$$

By definition,

$$H^{p_1}U(g) \leq H^{p_2}U(g).$$

From a given g, it is possible to compute $g'_a(\cdot)|_{p_1}$ for an arbitrary c and then the following will hold

$$V \leq U \Longrightarrow$$

$\forall g(w), c,$

$$V\left(g'_c(\cdot)|_{p_1}\right) \leq U\left(g'_c(\cdot)|_{p_1}\right) \Longrightarrow$$

$$\sum_{a\in A} V\left(g'_a(\cdot)|_{p_1}\right) \cdot p_1 g \leq \sum_{a\in A} U\left(g'_a(\cdot)|_{p_1}\right) \cdot p_1 g \Longrightarrow$$

$$\sum_{w\in W}\left(\sum_{a\in A} u(c,l)\,p_1\right) g(w) + \beta\sum_{a\in A} V\left(g'_a(\cdot)|_{p_1}\right) \cdot p_1 g$$
$$\leq \sum_{w\in W}\left(\sum_{a\in A} u(c,l)\,p_1\right) \Longrightarrow$$

$$H^{p_1}V(g) \leq H^{p_1}U(g) \Longrightarrow$$
$$H^{p_1}V(g) \leq H^{p_2}U(g) \Longrightarrow$$
$$HV(g) \leq HU(g) \Longrightarrow$$
$$HV \leq HU.$$

Note that g was chosen arbitrarily and, from it, $g'_a(\cdot)|_{p_1}$ completes the argument that the value function is isotone. ■

7.3 The Optimal Value Function is Piecewise Linear

Proposition 2. *If the utility is CRRA or LOG with a parameter of risk aversion* $\gamma \in (0, +\infty)$ *and inverse of Frisch elasticity of labor supply* $\eta \in [0, +\infty)$ *and if* Pr (a_j, w_i) satisfies (16)-(19), *then the optimal* $n-step$ *value function* $V_n(g)$ *can be expressed as:*

$$V_n(g) = \max_{\{\alpha_n^i\}_i} \sum_i \alpha_n(w_i)\, g(w_i)$$

where the $\alpha-vectors$, $\alpha : W \to R$, *are* $|W|-$*dimensional hyperplanes.*

Proof. The proof is done via induction. I assume that all the operations are well-defined in their corresponding spaces. Let Γ be the set that contains constraints (16)-(19) .For planning horizon $n = 0$, I have only to take into account the immediate expected rewards and thus I have that:

$$V_0(g) = \max_{p \in \Gamma} \left[\sum_{w \in W} \left(\sum_{a \in A} u(c,l)\, p \right) g(w) \right] \tag{21}$$

and therefore if I define the vectors

$$\left\{\alpha_0^i(w)\right\}_i \equiv \left(\sum_{a \in A} u(c,l)\, p \right)_{p \in \Gamma} \tag{22}$$

I have the desired

$$V_0(g) = \max_{\left\{\alpha_0^i(w)\right\}_i} \left\langle \alpha_0^i, g \right\rangle \tag{23}$$

where $\langle .,. \rangle$ denotes the inner product $\langle \alpha_0^i, g \rangle \equiv \sum_{w \in W} \alpha_0^i(w), g(w)$. For the general case, using equation (**??**):

$$V_n(g) = \max_{p \in \Gamma} \left[\begin{array}{c} \sum_{w \in W} \left(\sum_{a \in A} u(c,l)\, p(c,l|w) \right) g(w) + \\ +\beta \sum_{w \in W} \sum_{a \in A} \left(V_{n-1}\left(g'_a(\cdot)_a\right)\right) p(c,l|w)\, g(w) \end{array} \right] \tag{24}$$

by the induction hypothesis

$$V_{n-1}\left(g(\cdot)|_a\right) = \max_{\left\{\alpha_{n-1}^i\right\}_i} \left\langle \alpha_{n-1}^i, g'_a(\cdot) \right\rangle \tag{25}$$

Plugging into the above equation (**??**) and by definition of $\langle .,. \rangle$,

$$V_{n-1}\left(g'_a(\cdot)\right) = \max_{\left\{\alpha_{n-1}^i\right\}_i} \sum_{w' \in W} \alpha_{n-1}^i(w') \left(\sum_{w \in W} \sum_{a \in A} T(\cdot; w, c, l) \frac{\Pr(w,c,l)}{\Pr(c,l)} \right) \tag{26}$$

With the above:

$$V_n(g) = \max_{p\in\Gamma}\left[\begin{array}{c}\sum_{w\in W}\left(\sum_{a\in A} u(c,l)\,p\right) g(w) + \\ +\beta \max_{\left\{\alpha^i_{n-1}\right\}_i} \sum_{w'\in W} \alpha^i_{n-1}(w')\left(\sum_{w\in W}\left(\sum_{a\in A}\frac{T(\cdot;w,c,l)}{\Pr(c,l)}\cdot p\right) g(w)\right)\end{array}\right]$$

$$= \max_{p\in\Gamma}\left[\langle u(c,l)\cdot p,\ g(w)\rangle + \beta\sum_{a\in A}\frac{1}{\Pr(c,l)}\max_{\left\{\alpha^i_{n-1}\right\}_i}\left\langle \sum_{w'\in W}\alpha^i_{n-1}(w')\,T(\cdot;w,c,l)\cdot p,\ g\right\rangle\right] \tag{27}$$

At this point, it is possible to define

$$\alpha^j_{p,a}(w) = \sum_{w'\in W}\alpha^i_{n-1}(w')\,T(\cdot : w,c,l)\cdot p. \tag{28}$$

Note that these hyperplanes are independent on the prior g for which I am computing V_n. Thus, the value function amounts to

$$V_n(g) = \max_{p\in\Gamma}\left[\langle u(c,l)\cdot p,\ g\rangle + \beta\sum_{a\in A}\frac{1}{\Pr(c,l)}\max_{\left\{\alpha^j_{p,a}\right\}_j}\left\langle \alpha^j_{p,a}, g\right\rangle\right], \tag{29}$$

and define:

$$\alpha_{p,a,g} = \arg\max_{\left\{\alpha^j_{p,a}\right\}_j}\left\langle \alpha^j_{p,a}, g\right\rangle. \tag{30}$$

Note that $\alpha_{p,a,g}$ is a subset of $\alpha^j_{p,a}$ and using this subset results into

$$V_n(g) = \max_{p\in\Gamma}\left[\langle u(c,l)\cdot p,\ g\rangle + \beta\sum_{a\in A}\frac{1}{\Pr(c,l)}\langle \alpha_{p,a,g}, g\rangle\right]$$

$$= \max_{p\in\Gamma}\left\langle u(c,l)\cdot + \beta\sum_{a\in A}\frac{1}{\Pr(c,l)}\alpha_{p,a,g},\ g\right\rangle. \tag{31}$$

Now

$$\left\{\alpha^i_n\right\}_i = \bigcup_{\forall g}\left\{u(c,l)\cdot p + \beta\sum_{a\in A}\frac{1}{\Pr(c,l)}\alpha_{p,a,g}\right\}_{p\in\Gamma} \tag{32}$$

is a finite set of linear function parametrized in the action set. ∎

7.4 .. and Convex (PCWL)

Proposition 3. *Assuming the CRRA or LOG utility function and the conditions of Proposition 1, let V_0 be an initial value function that is piecewise linear and convex. Then the i^{th} value function obtained after a finite number of update steps for a rational inattention consumption-saving problem is also finite, piecewise linear and convex (PCWL).*

Proof. The first task is to prove that $\{\alpha_n^i\}_i$ sets are discrete for all n. The proof proceeds via induction. Assuming CRRA/LOG utility and since the optimal policy belongs to Γ, it is straightforward to see that through (22), the set of vectors $\{\alpha_0^i\}_i$,

$$\{\alpha_0^i\}_i \equiv \left(\sum_{w\in W}\left(\sum_{a\in A}\left(\frac{c^{1-\gamma}}{1-\gamma}-\alpha\frac{l^{1+\eta}}{1+\eta}\right)p\left(c,l|w\right)\right)g\left(w\right)\right)_{p\in\Gamma}$$

is discrete. For the general case, observe that for discrete controls and assuming $M = \left|\left\{\alpha_{n-1}^j\right\}\right|$, the sets $\left\{\alpha_{p,c}^j\right\}$ are discrete, for a given action p and consumption c, I can only generate $\alpha_{p,c}^j-$vectors. Now, fixing p it is possible to select one of the M $\alpha_{p,c}^j-$vectors for each one of the observed consumption c and, thus, $\{\alpha_n^j\}_i$ is a discrete set. The previous proposition, shows the value function to be convex. The *piecewise-linear* component of the properties comes from the fact that $\{\alpha_n^j\}_i$ set is of finite cardinality. It follows that V_n is defined as a finite set of linear functions. ■

8 Appendix B : Model Statistics and Graphs

8.1 Tables

$\theta = 2, \{\gamma = 1, \eta = 1\}$	Mean	St.Dev
Consumption (c)	3.55	1.79
Labor (l)	3.05	1.52
Wealth (w)	6.03	3.11
Information Flow (κ)	*0.73*	*0.81*

Table 10a: Statistical properties of the Model, M(2,1,1)

$\theta = 0.2, \{\gamma = 1, \eta = 1\}$	Mean	St.Dev
Consumption (c)	3.95	1.10
Labor (l)	2.68	1.06
Wealth (w)	5.84	1.86
Information Flow (κ)	*1.08*	*0.68*

Cross-Correlation

	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.72	1	0.60	0.72	0.66	0.68
(l)	0.74	0.72	0.54	1	0.81	0.74
(w)	0.78	0.68	0.63	0.74	0.61	1

Table 10b: Statistical properties of the Model, M(0.2,1,1)

$\theta = 0.02, \{\gamma = 1, \eta = 1\}$		
	Mean	St.Dev
Consumption (c)	4.16	0.95
Labor (l)	2.26	0.63
Wealth (w)	5.06	1.11
Information Flow (κ)	*1.52*	*0.42*

Table 10c: Statistical properties of the Model, M(0.02,1,1)

$\theta = 0.2, \{\gamma = 3, \eta = 1\}$		
	Mean	St.Dev
Consumption (c)	4.02	1.01
Labor (l)	2.90	0.96
Wealth (w)	5.90	1.74
Information Flow (κ)	*1.11*	*0.35*

Cross-Correlation						
	$c\,(-1)$	c	$l\,(-1)$	l	$w\,(-1)$	w
(c)	0.79	1	0.81	0.91	0.91	0.74
(l)	0.66	0.91	0.51	1	0.74	0.64
(w)	0.86	0.74	0.58	0.64	0.66	1

Table 10d: Statistical properties of the Model M(0.2,3,1)

$\theta = 0.2, \{\gamma = 1, \eta = 0.25\}$		
	Mean	St.Dev
Consumption (c)	4.11	0.69
Labor (l)	3.01	0.94
Wealth (w)	5.00	1.02
Information Flow (κ)	*1.21*	*0.29*

Cross-Correlation						
	$c\,(-1)$	c	$l\,(-1)$	l	$w\,(-1)$	w
(c)	0.92	1	0.59	0.73	0.66	0.71
(l)	0.54	0.73	0.63	1	0.82	0.89
(w)	0.68	0.71	0.77	0.89	0.72	1

Table 10e: Statistical properties of the Model M(0.2,1,0.25)

$\theta = 0.2, \{\gamma = 1, \eta = 4\}$		
	Mean	St.Dev
Consumption (c)	3.32	0.99
Labor (l)	2.20	0.74
Wealth (w)	5.07	1.31
Information Flow (κ)	*0.98*	*0.59*

Table 10f: Statistical properties of the Model M(0.2,1,4)

8.2 Figures

Comparing consumption for different utilities, θ=0.2

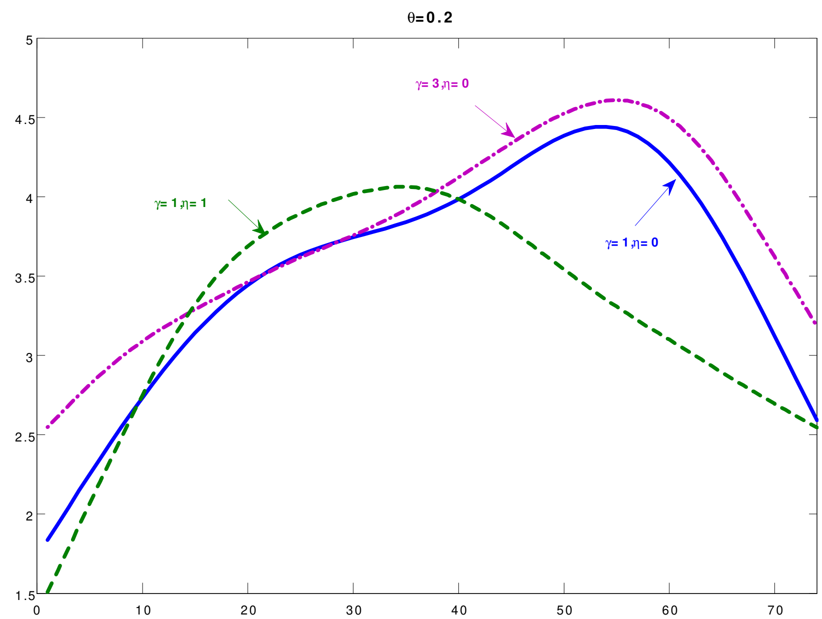

Figure 8 :Blue: γ=1,η=0; Green: γ=1,η=1; Violet:γ=3,η=0.

Comparing labor and wealth for different utilities, θ=0.2

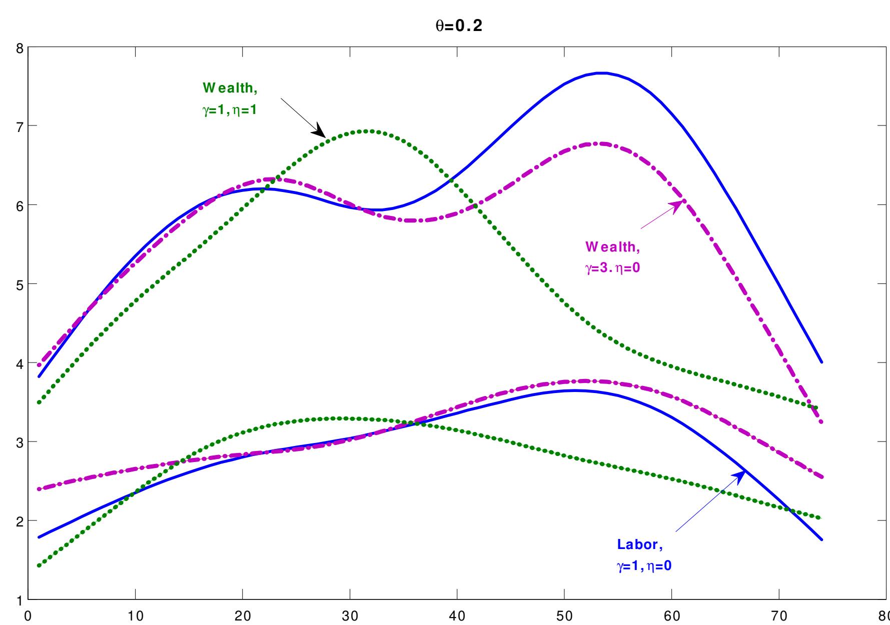

Figure 9 :Blue: γ=1,η=0; Green: γ=1,η=1; Violet:γ=3,η=0.

Comparing savings for different utilities, θ=0.2

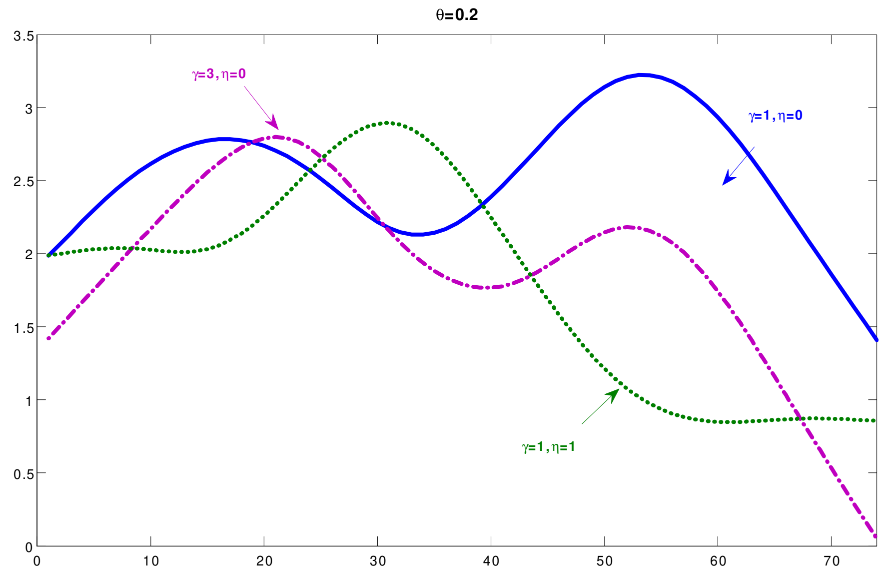

Figure 10 :Blue: γ=1,η=0; Green: γ=1,η=1; Violet:γ=3,η=0.

Comparing wealth, consumption, savings and income for different θ's, $u(c,l)=\log(c)-\alpha\frac{l^{1+\eta}}{1+\eta}$

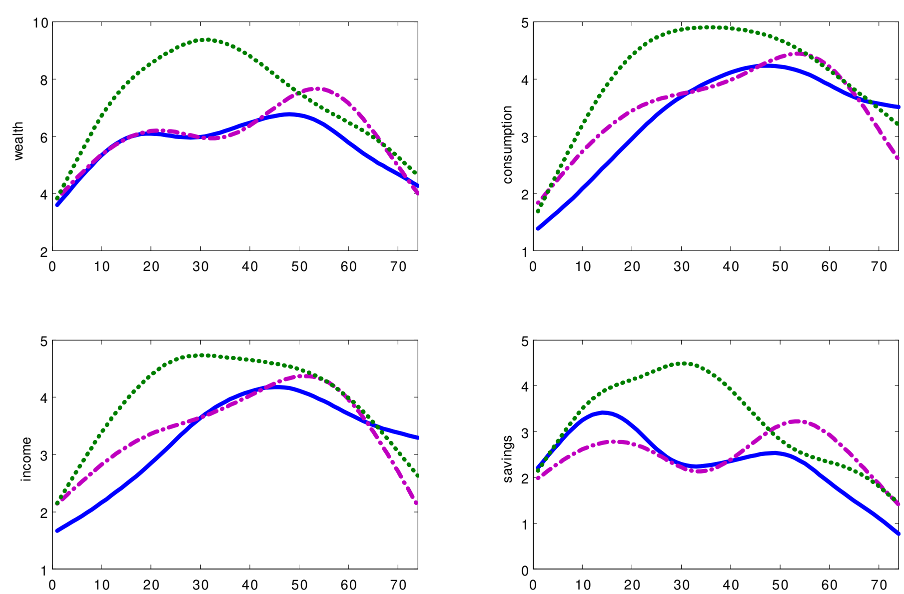

Figure 11 :Blue: θ=2; Violet:θ=0.2; Green: θ=0.02.

9 Appendix C: Rigidity of Labor and Consumption Choices

This section builds up a low-dimension intuition for the solution strategy of the model in the section 2 before turning to the formal solution and its findings. Consider a consumer who can choose to consume a quantity in the set $\Omega_c \equiv \begin{pmatrix} c_{low} & c_{high} \end{pmatrix}$. Each period, he decides whether to work $\Omega_l \equiv (\{l=0\} \vee \{l=1\})$ in exchange for a salary s. Assume for simplicity no asset but a fixed initial endowment $\bar{w}=2$. The budget constraint is:

$$c \leq \bar{w} + s\mathbf{1}_{\{l=1\}} \tag{33}$$

where $\mathbf{1}_{\{l=1\}}$ indicates whether the consumer works.

Let $u(c,l) \equiv \log c - \alpha l$ denotes the utility of the consumer. Moreover, to make matters concrete, let $c_{low}=2$, $c_{high}=4$ and $\alpha=0.3$ and $s=2$ with probability p and $s=1$ with probability $(1-p)$

Under full information capacity and no uncertainty, the agent will work iff:

$$\alpha c \leq s \tag{34}$$

The solution for this problem is clearly $(c_{high}, l=1)$ iff $p \leq 0.2$ and $(c_{low}, l=0)$ if.$p > 0.2$. Now assume that it is prohibitively costly for the agent to know the probability of the outcomes for s. In this case, it is optimal for the consumer to choose $(c_{low}, l=0)$.

Under rational inattention, the agent can reduce his uncertainty up to an amount given by his ability of processing information. Such a constraint, expressed in terms of change in entropy is the Shannon channel. The reduction in uncertainty is obtained by choosing the distribution of a signal informative about the underlying state (salary) as much as the Shannon channel allows it. In particular, there are 3 possible choices the consumer can make and that satisfy (33), i.e., $\left\{ \begin{array}{ccc} (c_{high}, l=1), & (c_{low}, l=1), & (c_{low}, l=0) \end{array} \right\}$. With the constraint that the joint distribution $p(\{c,l\},s)$ delivers as marginal for s $\Pr(s=1)=p$ and $\Pr(s=2)=1-p$, the joint distribution is

$C, L \backslash S$	s_1	s_2	
$(c_{low}, l=1)$	z_2	z_3	
$(c_{low}, l=0)$	$p - z_2$	$(1-p) - z_1 - z_3$	
$(c_{high}, l=1)$	0	z_1	

.

The problem of the consumer is then to

$$\begin{aligned} & \max_{z_i} E(u(c,l)) \\ = \; & u(c_{high}, l=1) z_1 + u(c_{low}, l=1)(z_2+z_3) + u(c_{low}, l=0)(1-z_1-z_2-z_3) \end{aligned}$$

s.t.

$$I(p(\cdot_s;\cdot_a)) = \sum_s \sum_{\{c,l\}} p(\{c,l\},s) \log\left(\frac{p(\{c,l\},s)}{\left(\sum_{s'} p(\{c,l\},s')\right) g(s)}\right)$$

The first order condition are

∂z_1 :

$$u\left(c_{high}, l=1\right) - u\left(c_{low}, l=0\right) = \lambda\left(\ln\left(\frac{z_1}{1-z_1-z_2-z_3}\right)\right)$$

∂z_2 :

$$u\left(c_{low}, l=1\right) - u\left(c_{low}, l=0\right) = \lambda\left(\ln\left(\frac{z_2+z_3}{z_1}\right) - \ln\left(\frac{1-z_1-z_2-z_3}{p-z_2}\right)\right)$$

∂z_3 :

$$u\left(c_{low}, l=1\right) - u\left(c_{low}, l=0\right) = \lambda\left(\ln\left(\frac{z_2+z_3}{z_3}\right) - \ln\left(\frac{1-z_1-z_2-z_3}{(1-p)-z_1-z_3}\right)\right)$$

The set of first order conditions yield a system of simultaneous trascendental equations. This system of transcendental equations involving logarithms can be solved using the LambertW function., which is an inverse mapping satisfying $W(y)\,e^{W(y)} = y$ and thus $\log W(y) + W(y) = \log y$. This function has multiple branches, Branches 0 and -1 are the only ones that can take on non-complex values. Let $y = e^x$. To solve the three equations, combine the last two F.O.C.'s and plug the solution in the first using the constraints on the marginals. Let $\omega_i \equiv u\left(c_i, l=1\right) - u\left(c_{low}, l=0\right)$ where $i \equiv \left(c_{high}, c_{low}\right)$. Then the solution for z_i is given by

$$z_i = \frac{-\lambda/\left(\omega_i+p\right)}{W\left(\left(\lambda/\left(\omega_i+p\right)\right)e^{-\lambda(1+p)/\omega}\right)}. \tag{35}$$

Figure 11 below illustrates the behavior of z_1 as function of λ, p and ω_1.

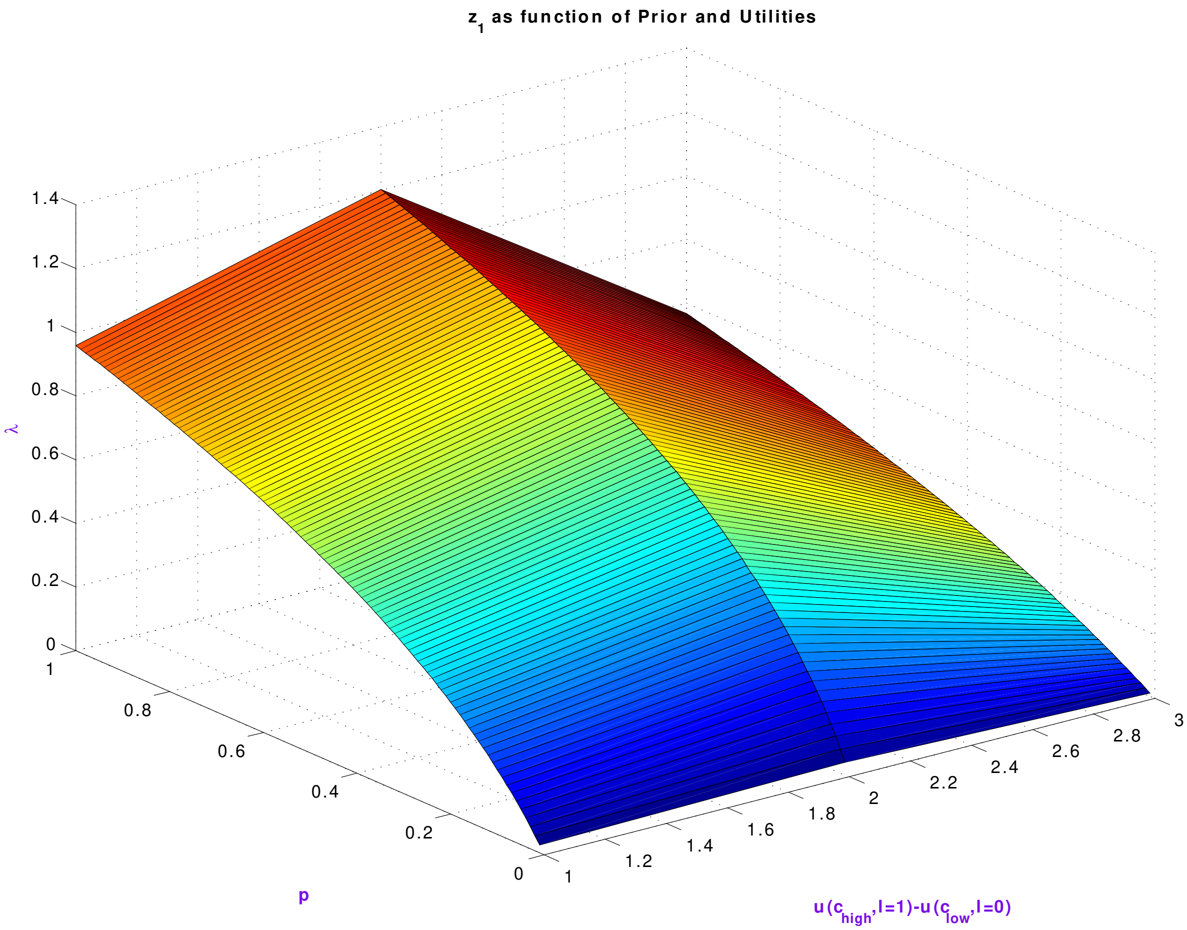

Figure 11 :An analitical solution

10 Appendix D

Pseudocode

Let θ be the shadow cost associated to $\kappa = I(A, W)$. Define a Model as a pair (γ, η, θ). For a given specification :

- Step 1: Build the simplex. Construct a uniform grid to approximate each $g(w)$-simplex point.

- Step 2: For each simplex point, define $p(w, c, l)$. and initialize with $V\left(g'_{a_j}(\cdot)\right) = 0$.

- Step 3: For each simplex point, find $p^*(w, a, c)$ s.t.

$$V_0(g(w))|_{p^*(w,c,l)} = \max_{p^*(w,c,l)} \left\{ \sum_{w \in W} \sum_{a \in A} \left(\frac{c_t^{1-\gamma}}{1-\gamma} - \alpha \frac{l^{\eta+1}}{\eta+1} \right)_{p^*(w,c,l)} - \theta[\kappa] \right\}.$$

- Step 4: For each simplex point, compute $g'_{a_j}(\cdot) = \sum_{w \in W} T(\cdot; w, c, a)\, p^*(w|c, l)$. Use a kernel regression to interpolate $V_0(g(w))$ into $g'_{a_j}(\cdot)$.

- Step 5: Optimize using csminwel and iterate on the value function up to convergence.

Obs. Convergence and Computation Time vary with the specification (γ, η, θ).

$\rightarrow$ 120-220 iterations each taking 8min-20min

- Step 6. For each model (γ, η, θ), draw from the ergodic $p^*(w, c, l)$ a sample (c_t, l_t, w_t) and simulate the time series of consumption, wealth, expected wealth and information flow by averaging over 1000 draws.

- Step 7. Generate histograms of consumption and impulse response function of consumption to temporary positive and negative shocks to income.