

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**The Rigidity of Labor: Processing Savings and Work Decisions
Through Shannon's Channels**

Antonella Tutino

2010-02

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

The Rigidity of Labor

Processing savings and work decisions through Shannon's channels

ANTONELLA TUTINO *

First version: May 2008 ; This version: February 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. My 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. My model also finds high persistence and strong comovement of consumption and employment and delayed response of consumption and labor with respect to wealth. Furthermore, my framework generates endogenously a wedge between marginal rate of substitution and marginal rate of transformation or wages. Such wedge is bigger and more volatile the lower information flow.

These findings suggest that rational inattention offers a promising avenue to bridge the gap between theory and U.S. business cycle data.

*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 aggregate behavior. Existing models match some U.S. business cycle facts at the expense of others. 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, starting from a framework in which consumers and firm optimize, market clears and no fiscal distortion, equilibrium in the labor market the marginal rate of substitution (MRS) between consumption and labor ought to be equal to the marginal product of labor (MPL) given by wages. As a matter of fact, the data shows a discrepancy between the two, called the labor wedge. The challenge for macroeconomic theories has been to match magnitude and volatility of such a wedge (Hall 1997).

Over the years, a number of explanations and modelling devices have been proposed to address the movement in the labor wedge over the business cycle. The principal suspects are different sorts of shocks to either the demand or the supply side of the labor market or both. Whether the shocks hit wage mark-ups,¹ disutility of labor,² taxes³ or productivity as in the search and matching literature,⁴ the common denominator seems to be the close connection between ad hoc modelling technologies and results. Moreover and most importantly, models based on exogenous shocks do not only fall short in micro-founded justification but also fail in providing empirically plausible predictions along other dimensions.⁵

This paper suggests an explanation of labor wedge and, more generally, business cycle facts, based on rational inattention.⁶ My choice to model people constrained by finite information processing capacity agrees with intuition and, as the paper shows, points towards the right direction to capture empirical evidence. In particular, my model is able to generate endogenously a wedge between MRS and MPL in an optimizing framework without the need of exogenous shocks.

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 disutility of labor. 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 they can attend to and decide on the basis of that information how much to consume and work. Realized consumption and employment are then 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 decision. Thus, fluctuations in wealth are due to

¹cfr. e.g., Galì, Gertler and Lopez-Salido (2007), Smets and Wouters (2007).

²cfr., e.g., Rotemberg and Woodford (1987), Smets and Wouters (2003), Galì and Rabanal (2004).

³cfr. e.g., Uhlig (2003), McGrattan and Prescott (2007).

⁴cfr. e.g., Pissarides (1985), Mortensen and Pissarides (1994), Shimer (1995), Hall (1995) and Galì and Trigari (2006).

⁵cfr. Pissarides (2007) and Shimer (2008).

⁶cfr., Shannon (1948), Sims (1988, 1998, 2003, 2006), Tutino (2008).

movements in labor and consumption as well as movements in wages. Consumers keep track of wealth by processing information in the limits of their Shannon capacity.

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 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 specificity of the 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. business cycle data along the dimensions analyzed. I show that even in a simple partial equilibrium 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 U.S. business cycle. In particular, my setting delivers four results: 1. 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 derived labor wedge with positive variance.

To understand the mechanism behind these results, consider what happens in an equivalent model with full information. When wages are deterministic and interest rates are high enough to promote savings, a consumer aware of his wealth with log utility and constant disutility of labor accumulates wealth through high labor supply which, in turns, increase period-by-period income while keeping consumption smooth. In such a context fluctuations in wealth are mostly due to fluctuations in labor supply. With information processing constraints, consumers cannot know the exact value of their wealth even with deterministic wages. 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 linear disutility of labor together with low information flow make household 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 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 disutility of labor. Such preferences imply that when people review their consumption plan by realizing that they have saved too much or too little with respect to their lifetime possibilities, they do so by large amounts. This 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 derive 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 implies that consumers are 'stuck' with the same choices 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. If information capacity is low, then contemporaneous consumption and labor commove 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 -optimal joint distribution between wealth and behavior. With high information flow, consumers' optimal policy commands to make wealth and behavior 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. These results are robust to higher specification of risk aversion and for the case of finite elasticity of substitution of leisure in different periods.

The last result is that my model derives endogenously a wedge between MRS and MPL. Moreover, a reasonable calibration of the model leads to magnitude and variance of this wedge in line with U.S. data ⁷. The key point is that people react to fluctuations of wealth according to the information they have chosen. As a result, consumption and labor are function of the optimal signal that consumers select. The lower the information flow, the less informative the signal, the more uncertain consumers are about their wealth. Consumers react to such uncertainty by modifying their consumption and labor on the basis of their perception of the wealth given the signal. This mechanism generates a discrepancy between the actual wage (MPL) and the marginal rate of substitution between consumption and labor (MRS) based on consumers' information.

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

⁷cfr. Shimer (2008).

a dynamic context and allows for risk aversion in the specification of agents' preferences. 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, and, in turn, a different degree of persistence between consumption and labor on the basis of the stochastic properties of the joint distribution of decisions and wealth chosen by the consumers. An example might help clarifying the intuition behind this result. 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 on both consumption and leisure. 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 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 business cycle data. Section 5 provides statistics and predictions of different specifications of the model. Section 6 concludes.

2 The Model

The model is a one sector partial equilibrium discrete time problem. The economy is populated by numerous households who want to maximize the expected discounted value of their utility defined over a consumption good and leisure. I assume that the consumers have infinite elasticity of substitution between leisure in different periods.⁸ Their wealth evolves according to their previous period savings augmented by a fixed and exogenous interest rate and their income is given by the hours of work they decide to supply in exchange for a salary.

The households cannot process all the information available about their wealth due to limited processing capacity. Instead, they have a belief about it. The assumption that

⁸The main results of the paper are not affected if one assumes finite elasticity of substitution. This is shown in Section 5.

wealth is not known is meant to capture the complexity associated with mapping current value of wealth (e.g., the number that people see at the end of their personal income statement) with their current and future possibility of consumption and leisure. The bound on information flow prevents households to perform this task immediately and with infinite precision. The model reflects households' initial uncertainty about wealth before processing any information through the assumption that people enter their life with a prior on wealth.⁹

Households decide which signals about wealth to choose in order to map their financial possibilities to consumption and work. They can freely choose the characteristics of their signals under the condition that the information between wealth and behavior provided by the signal flows at finite rate. 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.

2.1 Timing

Before the formal description, I provide a chronological view of the events to guide through the model. A person enters his life with a belief on wealth, $b(W)$. He decides on 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^*).

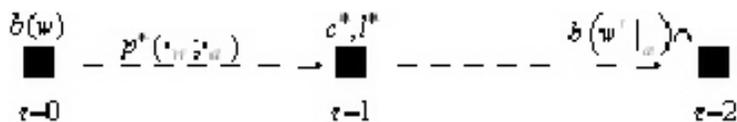


Figure 0: A typical day of a rationally inattentive person.

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.

⁹Cfr.: Sims (2003, 2006) and Tutino (2008).

2.2 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 consumers. Next I turn to the information-flow constraint, key of the model. Finally I present the objective function and set up the full problem.

The structure of the economy follows closely the one of Tutino (2008) to which I refers 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.

2.2.1 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 a budget constraint

$$W_{t+1} = R_t (W_t - C_t) + s_t L_t \quad (1)$$

where $R_t = R$ is the (constant) interest rate on savings, $(W_t - C_t)$, s_t is the wage the agent receive in exchange of L_t units of labor. The process characterizing the wages is Markovian with constant mean $E(s_t) = \bar{s}$. Consumers 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, $g(w_t)$, 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 at time t and similarly l_t is a specific outcome of the random variable L . Note that a is drawn from the optimal choice $p^*(\cdot_w; \cdot_a)$.

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

$$g'(w'|\tilde{a}) = \int_w T(w'; w, \tilde{a}) \cdot p(w|\tilde{a}) dw \quad (2)$$

where $T(w'; \cdot, \cdot)$ is the transition function commanded by the dynamic of wealth (1)¹⁰ and $p(w|\tilde{a})$ 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.

¹⁰Actually, the operator $T(w'; w, \tilde{a})$ assigns probabilities to w' conditional on the value of \tilde{a} and w . Since current values of w are not observable, the operator $T(\cdot)$ applies an expectation over the unknown

2.2.2 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 and 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.¹¹

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 $l^*(w)$. 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 values 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.

To understand how consumers' preference are necessary to determine $p(\cdot_w, \cdot_c, \cdot_l)$, consider a risk averse person. A risk averse individual would be more interested in knowing when his account is close to be overdrawn than a risk-neutral one. It follows that the risk averse type wants to process more information about low values of wealth which in turns implies low values of consumption and via the effect that income has on wealth (eq.1), also about low values of labor. If the consumer finds it extremely costly to process information, he watches out for those values of wealth, consumption and labor that reduce his utility the most by collecting information along this particular dimension. A consumer with relative better abilities of processing information can allocate his attention to news about several combinations of wealth, consumption and labor, ranked according to his utility.¹²

w . For a particular realization of $\tilde{a} = \{C = \tilde{c}, L = \tilde{l}\}$, the operators is defined as:

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

¹¹Alternatively, one can think of the choice of $p(\cdot_w, \cdot_c, \cdot_l)$ as equivalent for the consumers 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 skills.

¹²Exploring the interaction between information processing constraints and risk aversion is relatively novel to the literature in rational inattention which has focussed mainly on the Linear Quadratic Gaussian

From the above discussion, it follows that an optimal policy function for a rationally inattentive individual involves choosing $p(\cdot_w, \cdot_c, \cdot_l)$. Consumption and labor are drawn from the optimal policy $p^*(\cdot_w, \cdot_c, \cdot_l)$.

2.2.3 Information Constraint

As briefly 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 (Shannon (1954), Sims (2003, 2006)) 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 variable W provided by the observation of A (conditional 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. An intuitive analogy to Shannon capacity is the activity of a person's brain. Suppose the person reads about the drop in prices in the housing market. According to the time he has and his knowledge of real estate, his reading of the news can lead him to make significant steps towards buying an house right away. If the person does not have enough time to absorb the information about real estate and loan market and to map this information into his current and future net worth, he might want to push forward his decision to buy until he has figured out this link.

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 -mutual information between W and A -. 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.¹³ 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

(LQG) framework (Sims 2003, Mackowiak and Wiederholt (2008a, 2008b)). Cfr. Tutino (2008) for a discussion of the advantages of moving into a fully endogenous choice of signal with respect to the LQG framework.

¹³The approach of fixing the shadow cost of information instead of the capacity is adopted in Sims (2006) and Tutino(2008).

process information -i.e., low θ -, might find it optimal to keep close track of his wealth in order to enjoy combination of consumption and leisure that maximize his utility.

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

$$\kappa_t = \mathcal{I}_t(p(\cdot_w, \cdot_a)) = \int p(w_t, a_t) \log \left(\frac{p(w_t, a_t)}{(\int p(\tilde{w}_t, a_t) d\tilde{w}_t) g(w_t)} \right) dw_t da_t \quad (3)$$

The expression in (3) 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 same intuition gathered in the previous section, had the consumer had infinite processing capacity, he would be 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 makes the variables W and A independent of each other, $(\mathcal{I}(p(\cdot_w; \cdot_a)) \rightarrow 0)$. 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 person can process a finite amount of information, he attends to information that make his saving and labor decisions as close as possible to wealth.

2.2.4 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 (3) below. The control for their maximization is a signal $p(w, a)$ that solves : ¹⁴

$$V(g(w)) = \max_{p(w,a) \in \mathcal{D}} \sum_w \sum_a u(c, l) p(w, a) - \theta \kappa + \beta \left(\sum_w \sum_a (V(g'(w'|_a))) p(w, a) \right) \quad (4)$$

subject to

$$\kappa = \mathcal{I}(p(\cdot_w, \cdot_a)) = \sum_w \sum_a p(w, a) \log \left(\frac{p(w, a)}{(\sum_{\tilde{w}} p(\tilde{w}, a)) g(w)} \right) \quad (5)$$

and (2) and the requirement that $p(w, a) \in \mathcal{D}$ where \mathcal{D} is the set of all distribution that satisfies

$$\sum_a p(\tilde{w}, a) = g(\tilde{w}) \quad (6)$$

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

$$0 \leq p(\cdot_w; \cdot_a) \leq 1, \forall w, a \quad (7)$$

$$\sum_w \sum_a p(w, a) = 1. \quad (8)$$

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

$\partial p(w, a) :$

$$0 = u(c, l) + \beta V(\bullet|_a) + \theta \left(\log \left(\frac{p(w, a)}{\sum_{w''} p(w''; \cdot_a)} \right) - \sum_w \sum_a p(w, a) \frac{\partial (\sum_{w''} (w''; \cdot_a))}{\partial (p(w, a))} \right) \\ + \beta \left(\sum_w \sum_a \left[\frac{\partial (V(g'(w'|_{a_t})))}{\partial g(w'|_{a_t})} \frac{\partial g(w'|_a)}{\partial p(w, a)} \right] p(w, a) \right) \quad (9)$$

where

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

Define the solution to the optimization problem of the consumers as the distribution $p^*(\cdot_w; \cdot_a)$. Summing over all the possible realizations of W and C leads to the marginal distribution of labor

$$p^*(L) \equiv \sum_w \sum_{a \in \{A \setminus L\}} p^*(\cdot_w; \cdot_a) \quad (10)$$

and likewise summing over w and l , delivers the marginal distribution of consumption:

$$p^*(C) \equiv \sum_w \sum_{a \in \{A \setminus C\}} p^*(\cdot_w; \cdot_a) \quad (11)$$

where the notation $s \in \{S \setminus X\}$ indicate that the possible values s takes in S exclude the subset $X \subset S$. The realized outcomes $\{c_t^{\mathcal{I}}, l_t^{\mathcal{I}}\}$ ¹⁶ are then drawn from the optimal joint $p^*(\cdot_w; \cdot_c, \cdot_l)$. Appendix B derives the solution of a static, low-dimensional version of problem (4)-(8) that admits a quasi analytical solution.

¹⁵Note that the first order condition in (9) is valid for $\theta > 0$. If $\theta = 0$, then the probabilities $g(w)$ and $p(\cdot_w, \cdot_a)$ are degenerate. In this case, $\mathcal{I}(p(\cdot_w; \cdot_a)) = 0$, so using Fano's inequality (Thomas and Cover 1991),

$$c(\mathcal{I}(p(\cdot_w; \cdot_a))) = c(w) \\ l(\mathcal{I}(p(\cdot_w; \cdot_a))) = l(w)$$

which makes the first order conditions for this case the full information solutions.

¹⁶The superscript \mathcal{I} denotes the dependency of such outcome to the optimal distribution of uncertainty, $p^*(\cdot_w; \cdot_a)$.

3 Solution Strategy

To appreciate how consumers' preferences interact with the information processing constraint, I consider the following specifications for the utility function:

$$u(c, l) = \begin{cases} \frac{c^{1-\gamma}}{1-\gamma} - \alpha l \\ \log c - \alpha l \\ \log c - \frac{\alpha}{\eta+1} l^{\eta+1} \end{cases} .$$

where γ is the coefficient of risk aversion on the utility of consumption, η is the inverse of Frisch elasticity of labor supply, ε , and $\alpha \in [0, 1]$ is a constant disutility associated to labor.

The computation methodology follows closely Tutino(2008) to which I refer for technical details. However, there is a computational difference that I shall highlight. The difference concerns the construction of the simplex. 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 which belong to the interval $[0, 1]$ and are subject to the condition that their sum is 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 (2) using a kernel regression of $V(\cdot)$ into $g'(w'|_a)$. I use an Epanechnikov kernel with smoothing parameter $h = 3$.¹⁷ Table 1 reports the benchmark

¹⁷The reason behind $h = 3$ comes from experimental trials for $h \in [2, 5]$ with increments of 0.5 (*i.e.*, $h = 2, h = 2.5, \dots$). While the results do not change substantially as I vary h in this range, the computational time is lower when I set $h = 3$.

parameter values and the grids.¹⁸ The value iteration converges in about 180 iterations.

Benchmark Values

	Discretization
Core State Space W	$[1 : 1 : 10]$
Consumption Space C	$\frac{8}{15} [1 : 1 : 10]$
Labor Space L	$[1 : 0.5556 : 6]$
Joint Distribution per simplex point	100×10
Marginal C	10×1
Marginal L	10×1
Marginal W	10×1
\bar{s}	1.2
α	0.02
γ	3
η	0
R	1.03
Discount Factor, β	0.9966

Table 1

I shall point out another caveat for the model calibration, concerning the ratio of the interest rate, R , and the inverse of the discount factor, β . The calibration in Table 1 implies that the annualized rate of return induces consumers to save until wealth reaches its maximum. Under full information and with utility $u(c, l) = \log(c) - \alpha * l$, such a high interest rate associated to constant disutility of labor makes the consumer accumulate wealth by varying labor supply more than consumption. Under rational inattention, infrequency and sharpness of the adjustments of consumption and labor makes the variance of consumption higher than the one of labor. To highlight the sharp differences between full information models and rational inattention ones, I maintain the assumption $1 < R\beta$ throughout the numerical simulations. The parametrization of the model is summarized in Table 1.

¹⁸The strategy used to reduce the dimensionality of the choice space is to pair $\{c_i, l_j\}$ per each row and associate to them a value of wealth. In a 2 states, 2 choices of l and 2 choices of c case, that results into:

$$\begin{array}{rcc}
 & w_1 & w_2 \\
 \{c_1, l_1\} & p_1 & p_2 \\
 \{c_2, l_1\} & p_3 & p_4 \\
 \{c_1, l_2\} & p_5 & p_6 \\
 \{c_2, l_2\} & p_7 & p_8
 \end{array}$$

The corresponding marginals are then evaluated as e.g., $p(l_1) = (p_1 + p_2) + (p_3 + p_4)$.

4 Predictions of the Model and the Business Cycle.

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. All the series have quarterly frequency from 1964.I to 2007.II.. I construct figures and statistics by detrending the data with the HP filter, using a value of $\lambda = 1600$. I then used the detrended data to compute mean, standard deviation and the 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, 1999) and Sims (1998, 2003)). The choice of HP filter over other filter 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 $\eta = 0$. 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 ($\theta = 2$). As for these particular numerical values, I verify empirically that given the discretization of core states and core decisions and the baseline model with log utility of consumption and linear 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 costs. 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, for values of θ below 0.05 deliver a solution very close to the full information case.

To get a quantitative assessment of what these costs mean, I compute the average difference in lifetime utility between the infinite capacity case and each of the θ -cases considered. That is, I set $u(c, l) = \log -\alpha l$ and I compute $E(u(c, l) - u(c^{\mathcal{I}}, l^{\mathcal{I}}, \theta))$ where $E(u(c, l))$ is average lifetime utility under infinite capacity case while $E(u(c^{\mathcal{I}}, l^{\mathcal{I}}, \theta))$ is the utility under a particular value of θ when θ takes values in $(\{2\}, \{0.2\}, \{0.02\})$. With an average value of $E(u(c, l)) \simeq 1.5$ under full information, $\theta = 2$ is associated to a loss of about 20% in lifetime utility, $\theta = 0.2$ implies a loss of about 11% while $\theta = 0.02$ delivers a loss of about 8%. The values used for this computation can be found in Table 10a – 10c in appendix B. I assume that an economy is populated evenly by those three types.¹⁹

The computations are as follows. For each θ , mean, standard deviation and correlation

¹⁹As it is possible to see from the statistics for each θ in Table 10a-10e 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 .

of the simulated series are calculated after I take averages of 7,000 Monte Carlo runs and 10! simplex points. The statistics for each of this series are in Tables 10a – 10c. In tables 10d – 10e the same methodology produces the results for the cases $M = \{0.2, 3, 0\}$ and $M = \{0.2, 1, 1\}$ respectively. 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 my 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 *RC* which stands for Rigidity of Choice.

4.1 Business Cycle Facts through Shannon’s lenses.

I focus on a set of business cycle facts regarding consumption, labor and their volatility, persistence and comovement. Then I derive the labor wedge²⁰ that the model produces. Finally I calibrate the model to match the first moment of the labor wedge from the data and see what costs of information processing it implies.

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

My model predicts excess volatility of consumption with respect to labor supply. This results 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 constant disutility of labor. 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 informativeness 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 $M(\theta, 1, 0) \equiv \frac{1}{3}(M(2, 1, 0) + M(0.2, 1, 0) + M(0.02, 1, 0))$

²⁰cfr. Prescott (2004) and Shimer (2005).

to the 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)	1.13
Labor (l)	0.98

Table 2: Statistical properties of US Business Cycle and Model

The explanation that my model provides to BC 1 can be summarized by the following result:

RC Finding 1. *For the model $M(\theta, 1, 0)$, the standard deviation of consumption over labor is 1.15. The volatility is higher the higher the information costs. For a given degree of risk aversion, higher information costs make consumption and labor more sluggish with sudden adjustments following accumulation of wealth. This behavior of consumption and labor generates higher variance.*

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

In my 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.

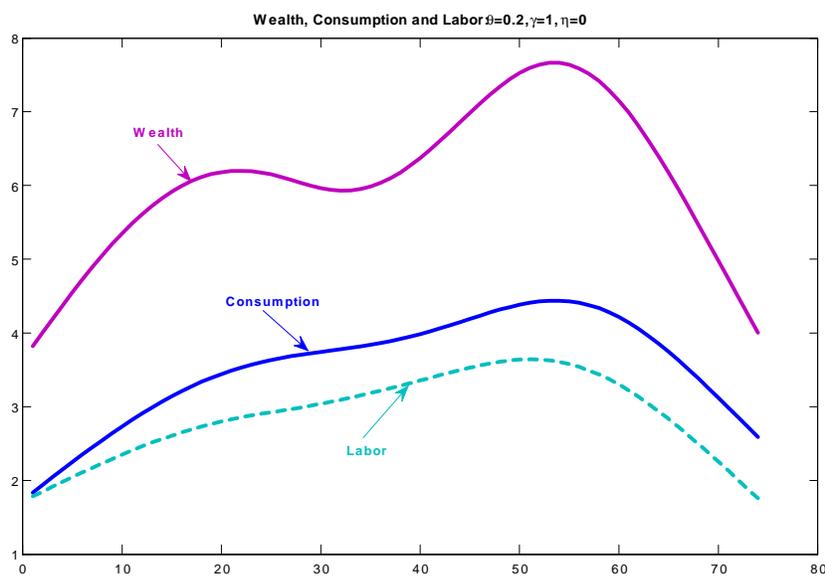


Figure 1: Sample Path of Macroeconomic variables, Average across 10,000 monte Carlo, $M=\{0.2,1,0\}$.

The intuition for these findings lies on the mechanism through which consumers update their knowledge of wealth, expressed in (2). 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 which trigger a reaction in behavior. This process is reflected in both delayed response to fluctuations of wealth and strong autocorrelation between current and past values of consumption and labor.

The comparison between models and findings is in Table 3.

US Business Cycle : Quarterly Data (1964.I-2007.II), HP filter, %		
	Std.Dev	Autocorr
Non Durable Goods (C)	0.84	93
Average Hours (L)	0.54	88
Simulated Data, average across models		
	Std.Dev	Autocorr
Consumption (c)	1.13	92
Labor (l)	0.98	87

Table 3: Statistical properties of US Business Cycle and Model

The explanation for BC Fact 2 can be summarized as:

RC Finding 2. *For the model $M(\theta, 1, 0)$, the autocorrelation of consumption and labor are above 80%. Moreover, contemporaneous consumption and labor lag wealth by one period.*

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

My 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 my model is twofold. First, the variables are related through the budget constraint (1) which is used by the consumers to update their prior on wealth. Second, the optimal policy of the consumer is the stationary joint distribution between choices (consumption and labor) and state (wealth).

Cross-Correlation		
	$w(-1)$	w
c	0.71	0.61
l	0.74	0.69
w	0.91	1

Table 5: Cross correlation of wealth; average across $M\{\theta, \gamma=1, \eta=1\}$

For a given shadow cost and disutility of labor, the higher the coefficient of risk aversion the higher the correlation between lagged and contemporaneous consumption and wealth. The same is true for labor and wealth. This results arises from the observation that risk averse individuals, say, $\gamma = 3$ pay more attention to wealth than individuals with $\gamma = 1$ for a given cost of processing. Controlling for the cost of information and the degree of risk aversion, people with increasing disutility of labor, $\eta = 1$, pay more attention to wealth than people with $\eta = 0$. However, this higher information is used to both consumption smoothing and increasing leisure. Even though the substitution effect prevails over the income effect also for $\eta = 1$, the resulting impact on the correlation between wealth and consumption is lower than for the case $\eta = 0$.

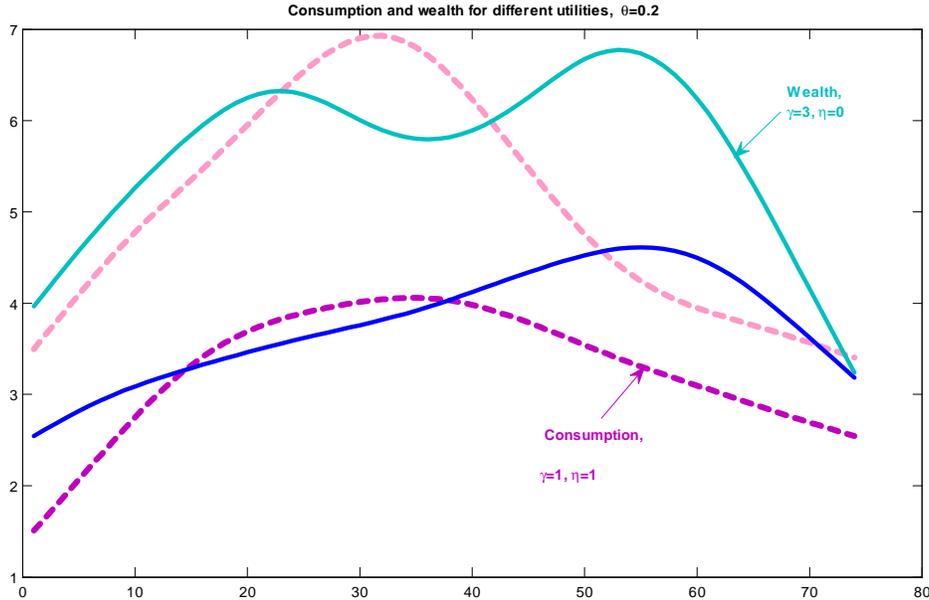


Figure 2: Consumption and wealth, $\theta=0.2$. Solid line: $\gamma=3, \eta=0$; Dotted Line: $\gamma=1, \eta=1$.

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 information costs. High coefficient of risk aversion 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 .

Simulated Data: Cross-Correlation				
	$c(-1)$	c	$l(-1)$	l
c	0.90	1	0.83	0.90
l	0.71	0.90	0.84	1

Table 6: Cross Correlation of US business cycle and averages across models

The explanation of the third business cycle fact from the model is:

RC Finding 3. *For the model $M(\theta, 1, 0)$ the contemporaneous cross correlation between consumption and labor is 0.90%.*

4.1.4 BC Fact 4: U.S. labor wedge has mean 0.4 and standard deviation $\cong 0.11$

US Business Cycle : Labor wedge 1994-2006, %		
	mean	std. dev.
Ψ	0.40	0.11

source: Shimer (2008)

Table7: Shimer's BC Fact 4

In absence of information-processing frictions ($\theta = 0$), and assuming an household with utility $u(c, l) = \log(c) - \tilde{\alpha} * l$ the labor wedge for the model is²¹

$$\Psi \equiv s - \tilde{\alpha} * c \tag{12}$$

where the wage, s , represents the marginal productivity of labor (MPL) and $U_l/U_c = -\tilde{\alpha} * c$ is the marginal rate of substitution (MRS). Prescott (2004) and Shimer (2008) estimates the average wedge Ψ to be 0.4 in U.S.

When information flows at finite rate ($\theta > 0$), consumers never observe wages directly but only through the impact wages have on signals about wealth. Thus, the MRS for information-constrained consumers has to be equal to their perception of the wages given the signal, $p^*(\cdot_w; \cdot_a)$. Define such perception of wages $\tilde{s}(\mathcal{I}(p(\cdot_w; \cdot_a)))$. Then, the wedge in my model is defined as $\nabla \equiv |\tilde{s}(\mathcal{I}(p(\cdot_w; \cdot_a))) - s|$ or:

$$\nabla = \underbrace{s}_{MPL} - \underbrace{\tilde{\alpha} * c(\mathcal{I}(p^*(\cdot_w; \cdot_a)))}_{MRS} \tag{13}$$

which increases with θ .

²¹cfr. e.g., Rotemberg and Woodford (1991, 1999), Hall(1997), Shimer (2008).

To get a sense of the magnitude involved, to obtain an average wedge $\nabla \simeq 0.4$, I set $\tilde{\alpha} = 0.24$ and $\theta = 2$. This specification implies the statistics in Table 8.

$\theta = 2, \tilde{\alpha} = 0.24$			
	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	3.20	1.08	0.93
Labor (l)	2.28	0.98	0.89
Wealth (w)	6.04	1.19	0.74
Information Flow (κ)	0.93	0.71	

Table 8: Sample model statistics. Average 7,000 Monte Carlo

In particular, the average consumption is $c(\mathcal{I}(p^*(\cdot_w; \cdot_a))) = 3.20$ and the consumption loss from not being well informed is about 16% of households' average lifetime consumption.²²

Even though the model cannot address jointly the supply and demand sides of the labor market²³ and it is limited to the intensive margin, I think it offers a promising angle to explain the discrepancy between the marginal rate of substitution and the marginal product of labor. Several studies have been trying to match theory and data to explain this wedge²⁴: cyclical fluctuation of hours driven by variation in taxes²⁵, time varying disutility of labor and worker's market power²⁶ and search and matching models²⁷.

Rather than examining the origins and bargaining of wage contracts, the suggestion of my model is to look more closely at consumers' behavior. In particular, my results show that adding a bound on information processing generates endogenously a wedge between the wage that prevails on the market and the perception of the workers of such wages, $\nabla \equiv |\tilde{s}(\mathcal{I}(p(\cdot_w; \cdot_a))) - s|$. In my model this labor wedge ∇ depends on the information-processing capacity of the consumers. Furthermore, the variance of ∇ depends on their choice of the informativeness of the signal. To see why, suppose that a person knows that he is rich in a given period of his life. Since information-processing is costly, he may decide to care a little about his current wealth and, as a result, $MRS > s$. By contrast, a person who doesn't have a lot of wealth wants to pay more attention to his financial activities which leads to a better monitoring of the wealth, or $MRS < s$. Note also that the choice of processing different flows of information according to consumers' needs and preferences make the variance of $\tilde{s}(\mathcal{I}(p(\cdot_w; \cdot_a)))$ positive.

²²To see why, I compute the loss as the information rent, $E(c^* - c(\mathcal{I}(p^*(\cdot_w; \cdot_a))))$, where $E(c^*)$ is the average consumption path when $\theta = 0$ which is equal to 3.78 and the expectation $E(\cdot)$ is taken with respect to time. This means that the cost of being uninformed is about 16% of household's average lifetime consumption.

²³cfr. Gertler and Trigari (2006).

²⁴cfr. Shimer (2008) and Pissarides (2007).

²⁵cfr. Uhlig (2003), Chen, Imrohrouglu and Imrohrouglu (2007).

²⁶cfr. Smets and Wouters (2007), Galí, Gertler and Lopez-Salido (2007).

²⁷cfr. Pissarides (1994), Shimer (2005), Hall(2005)

US Business Cycle : labor wedge 1994-2006, %

	mean	std. dev.
Ψ	0.40	0.11
Simulated data: labor wedge		
	mean	std. dev.
∇	0.40	0.41

Table 9: Labor wedge in data and model

This last result are summarized by the last findings.

RC Finding 4. *The labor wedge has mean 0.40 and standard deviation 0.41. The statistical properties of the wedge derive from those of the endogenously chosen distribution $p^*(\cdot_w; \cdot_a)$.*

5 A closer look at Model Statistics

The results of the previous section are robust to different characterizations of $M(\theta)$. Once the value function for each model converged, I simulate the path of consumption, labor, wealth and information flow for $T = 75$ averaging the results along the dimension of the simplex as well as 7,000 different Monte Carlo runs.

To get a sense on how the shadow cost of information affects consumption and labor behavior, consider $M(\theta) \equiv \{\theta; (\gamma \rightarrow 1, \eta = 3)\}$ where $\theta = 2, 0.2, 0.02$. Table 10a – 10e shows the statistics for these and $\gamma = 3, \eta = 1$ cases. All tables are collected in Appendix B.

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 the life-cycle. Moreover, with constant disutility of labor, ($\eta = 0$) labor supply adjusts according to wealth fluctuations to accommodate consumption smoothing. When information flows at finite rate, rational households choose signal about wealth with the same purpose. If processing-information has low cost, $\theta = 0.02$, consumers can choose a signal about wealth so informative 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, do not modify consumption and labor often. 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 θ -scenarios is in Figure 3.

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

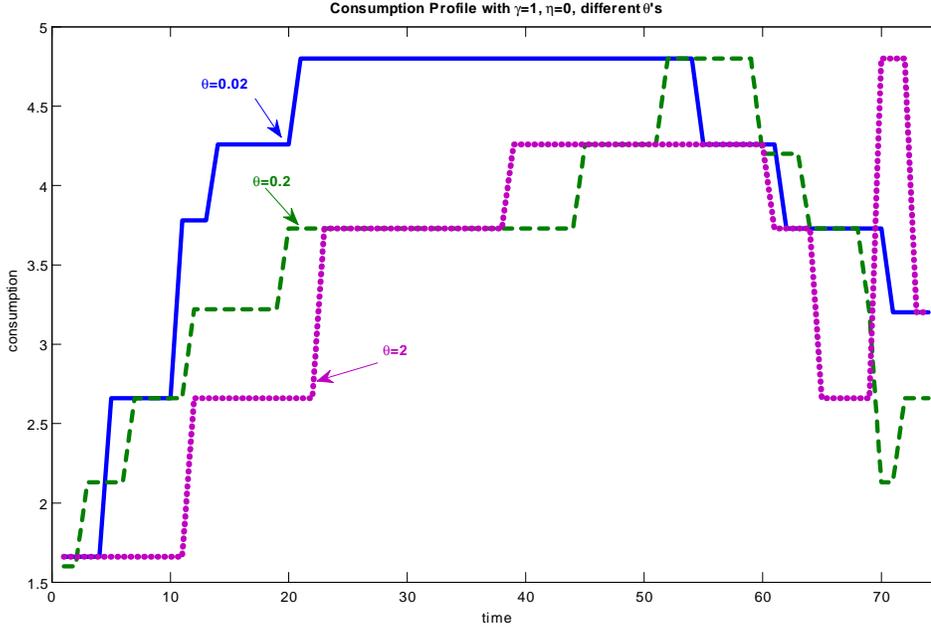


Figure 3: Sample path of consumption for different θ 's.

Figure 4 confirms the intuition that consumption is smoother the lower information costs. 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 results into slow and consistent 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 forwards an increase in consumption more than the other people ($\theta = 0.2$ and $\theta = 0.02$). 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 good purchases, though they manage to

enjoy some leisure at the end of the simulation (see Figure 5).

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

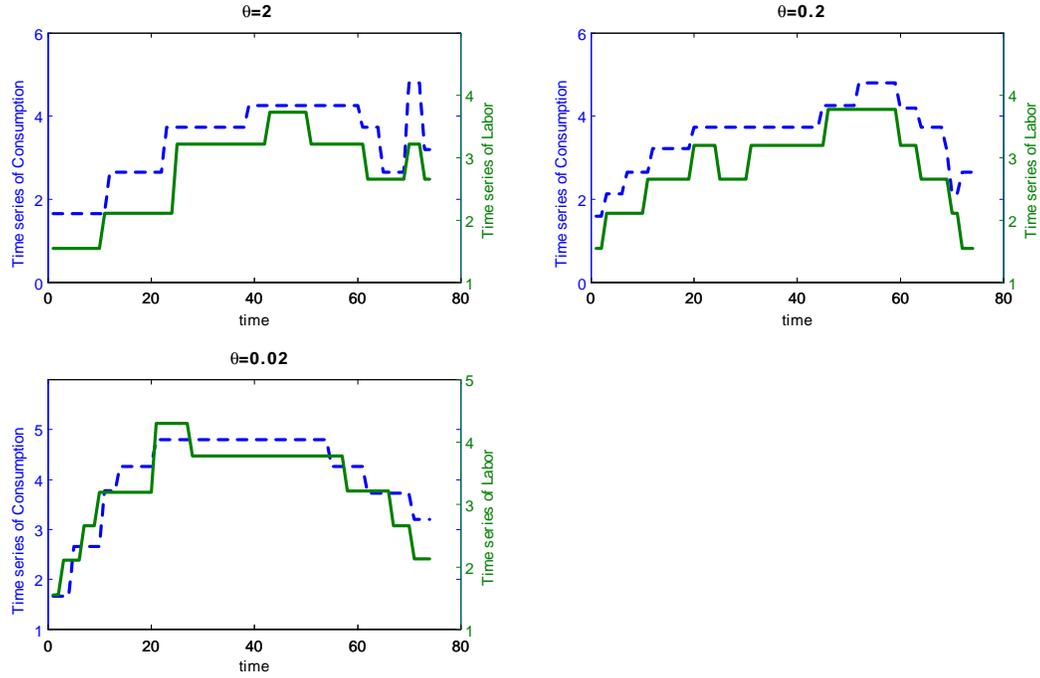


Figure 4: Time path consumption and labor, various θ

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.

Comparing consumption(blue) and wealth (red) for $u(c,l)=\log(c)-\alpha l$

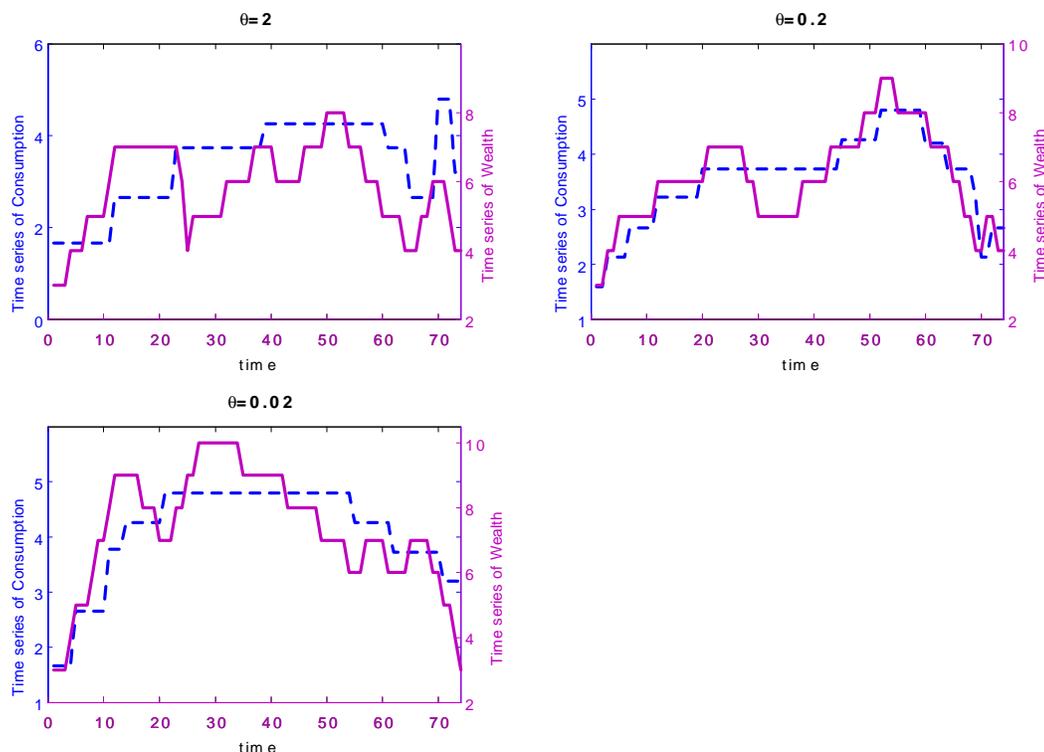


Figure 5: Time path wealth and consumption, log-lin utility, various θ

People who are more constrained in their choice of the signal, adjust with delays consumption to fluctuations in wealth. Such delays smooth consumption while consumers are processing information but at the same time, calls 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 3-5 illustrate also the high persistence of the series documented in table 12a – 12d. Not surprisingly the persistence is higher the higher the information cost .

For completeness, it is worth mentioning how the model behaves with higher coefficient of risk aversion and positive and finite elasticity of substitution. Figure 6 displays the stationary marginal distribution of consumption and wealth when $\gamma = 1$ and $\eta = 0$. These marginal distributions are computed from the joint distribution to which the value iteration converges as average over initial beliefs about wealth.

For low values of θ , households choose to be better informed about extreme values of consumption so that they can enjoy a balanced consumption throughout their lifecycle. This is, in essence, consumption smoothing.

Comparing marginal distributions of consumption and labor when $u(c,l)=\log(c)-\alpha l$

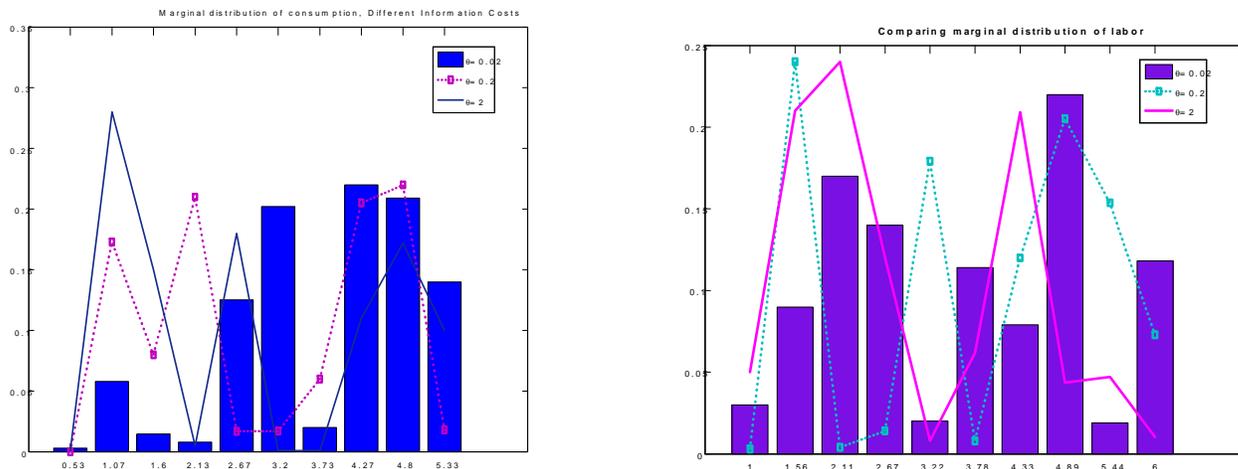


Figure 6: Marginal Distribution of consumption and labor various θ

The optimal marginal probability of labor assigns high probability on low labor supply when θ is low: as the capacity increases, consumers want to smooth consumption and sacrifice their leisure to do so (recall, labor effort has constant disutility in this case, $\eta = 0$). If information is costly to process, consumers will keep constant their consumption and labor effort. This in turn implies working less

When the signal about wealth is less informative on a period-by-period basis (high θ), income effect takes over the substitution effect: consumers are less certain that their wealth is high enough to enjoy leisure and as a result increase the keep constant their consumption and labor effort. The marginal probability distribution of labor reflects this pattern by allocating higher probabilities on lower values of labor (and consumption) the higher θ is.

Tables 10d – 10e show the relationship between risk aversion, γ , Frisch elasticity of substitution, $1/\eta$, and shadow cost of information, θ . 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 throughout later on due to the accumulated savings. Savings come from both low consumption and high labor supply at the beginning of the simulation 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 occur later in the simulation while labor supply is higher at the beginning than it is later on. Thus, information costs enhance precautionary savings.

Finally, for a given θ and γ , a lower the Frisch elasticity of substitution (from $\eta = 0$ to $\eta = 1$) 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 increase 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 case of perfect information and $\eta = 1$, 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 decreased. The opposite occurs when consumers receive more and more information about a decrease in wealth.

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: 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 and an endogenous labor wedge. In particular, the last prediction has been a challenge for macroeconomists. Existing macroeconomic theories have been explaining the labor wedge using exogenous shocks and ad-hoc market frictions, without micro-founded justifications and, more importantly, without a correspondence in the data. In my setting the difference between marginal rate of substitution and marginal product of labor, i.e., wage, occurs as a result of 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. This mechanisms results in a discrepancy between the marginal rate of substitution of consumption and labor and wages. Furthermore, the magnitude and the variance of the endogenous wedge are in line with the data. The same mechanism is able to capture persistence and lags of the main macroeconomic variables over U.S. business cycle. The findings of this paper suggest that making a leap to a rational inattention setting is worth the computational effort. For it gets us closer to understand and interpret empirical regularities in 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 am grateful to Per*

Krusell for his insightful advice and stimulating discussions. 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 Galí, 2006. “A New Keynesian Model with Unemployment.” MIT Mimeo.
- [5] Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, 2007. “Business Cycle Accounting.” *Econometrica*. 75 (3): 781–836.
- [6] 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.
- [7] 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.
- [8] Galí, Jordi, Mark Gertler, and J. David Lopez-Salido, 2007. “Markups, Gaps, and the Welfare Costs of Business Fluctuations.” *The Review of Economics and Statistics*. 89 (1): 44–59.
- [9] Galí, 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.
- [10] Gertler, Mark, and Antonella Trigari, 2006. “Unemployment Fluctuations With Staggered Nash Wage Bargaining.” NBER Working Paper 12498.
- [11] 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
- [12] Hall, Robert E., 1997. “Macroeconomic Fluctuations and the Allocation of Time.” *Journal of Labor Economics*. 15 (1): 223–250.
- [13] Hall, Robert E., 2005. “Employment Fluctuations with Equilibrium Wage Stickiness.” *American Economic Review*. 95 (1): 50–65.
- [14] Hansen, Gary D., 1985. “Indivisible Labor and the Business Cycle.” *Journal of Monetary Economics*. 16 (3): 309–327.

- [15] Imai, Susumu, and Michael P. Keane, 2004. “Intertemporal Labor Supply and Human Capital Accumulation.” *International Economic Review*. 45 (2): 601–641.
- [16] Kaldor, Nicholas, 1957. “A Model of Economic Growth.” *The Economic Journal*. 67 (268): 591–624.
- [17] 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.
- [18] Kalai, Gil. Linear programming, the simplex algorithm and simple polytopes. *Math. Programming*, 79(1-3,Ser. B):217–233, 1997.
- [19] Krusell, Per, and Anthony A. Smith, Jr., 1998. “Income and Wealth Heterogeneity in the Macroeconomy.” *Journal of Political Economy*. 106 (5): 867.
- [20] Lovejoy, W. Computationally feasible bound for partially observable Markov Decision problem, *Operations Research*, 39:162:175,
- [21] Lucas, Robert E. Jr., and Leonard A. Rapping, 1969. “Real Wages, Employment, and Inflation.” *The Journal of Political Economy*. 77 (5): 721–754.
- [22] Mackowiak, B., and M. Wiederholt, 2008a, "Optimal Sticky Prices under Rational Inattention", *American Economic Review*, forthcoming
- [23] Mackowiak, B., and M. Wiederholt, 2008b, "Business Cycle Dynamics under Rational Inattention", Mimeo.
- [24] MaCurdy, Thomas E., 1981. “An Empirical Model of Labor Supply in a Life-Cycle Setting.” *Journal of Political Economy*. 89 (6): 1059–1085.
- [25] Mankiw, N. Gregory, 1989. “Real Business Cycles: A New Keynesian Perspective.” *Journal of Economic Perspectives*. 3 (3): 79–90.
- [26] McGrattan, Ellen R., and Edward C. Prescott, 2007. “Technical Appendix: Unmeasured Investment and the Puzzling U.S. Boom in the 1990s.” *Federal Reserve Bank of Minneapolis Staff Report 395*.
- [27] 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*.
- [28] Modigliani, Franco, 1977. “The Monetarist Controversy or, Should We Forsake Stabilization Policies?.” *American Economic Review*. 67 (2): 1–19.
- [29] 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.

- [30] Mulligan, Casey B., 2002. "A Century of Labor-Leisure Distortions." NBER Working Paper 8774.
- [31] Parkin, Michael, 1988. "A Method for Determining Whether Parameters in Aggregative Models Are Structural." Carnegie-Rochester Conference Series on Public Policy. 29: 215–252.
- [32] Pissarides, Christopher A., 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *The American Economic Review*. 75 (4): 676–690.
- [33] Pissarides, Christopher A., 2007. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?." LSE Mimeo.
- [34] Prescott, Edward C., 2004. "Why Do Americans Work So Much More Than Europeans?." *Federal Reserve Bank of Minneapolis Quarterly Review*. 28 (1): 2–13.
- [35] Ragan, Kelly S., 2006. "Taxes, Transfers, and Time Use: Fiscal Policy in a Household Production Model." Ph.D. thesis, University of Chicago.
- [36] Rogerson, Richard, 1988. "Indivisible Labor, Lotteries and Eequilibrium." *Journal of Monetary Economics*. 21 (1): 3–16.
- [37] Rogerson, Richard, 2007. "Taxation and Market Work: Is Scandinavia an Outlier?." *Economic Theory*. 32 (1): 59–85.
- [38] Rogerson, Richard, and Johanna Wallenius, 2007. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." NBER Working Paper 13017.
- [39] Romer, Christina D., and David H. Romer, 2007. "A Narrative Analysis of Postwar Tax Changes." University of California, Berkeley Mimeo.
- [40] Rotemberg, Julio J., and Michael Woodford, 1991. "Markups and the Business Cycle." *NBER Macroeconomics Annual*. pp. 63–129.
- [41] Rotemberg, Julio J., and Michael Woodford, 1997. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual*. 12: 297–346.
- [42] Sims, Christopher A., 1998, "Stickiness, Carnegie-rochester Conference Series On Public Policy, 49(1), 317–356.
- [43] Sims, Christopher A., 2003, "Implications of Rational Inattention", *Journal of Monetary Economics*, 50(3), 665–690.
- [44] Sims, Christopher A., 2006 "Rational Inattention: A research Agenda", *American Economic Review Paper and Proceedings*.
- [45] Tutino, Antonella, 2008 "The Rigidity of Choices, Lifetime savings under information processing constraints". Princeton Universtiy, Mimeo.

- [46] Uhlig, Harald, 2003. "How Well do we Understand Business Cycles and Growth? Examining the Data with a Real Business Cycle Model." in Wolfgang Franz, Hans Jürgen. Ramser, and Manfred Stadler (ed.), *Empirische Wirtschaftsforschung: Methoden und Anwendungen*, vol. 32 of *Wirtschaftswissenschaftliches Seminar Ottobeuren*, pp. 295–319.
- [47] S. Verdú, 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.
- [48] S. Verdú, 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 $\|\cdot\|$ 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

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

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

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

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,

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

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

$$V \leq U \implies HV \leq HU$$

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

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

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

$$V \leq U \implies$$

$\forall g(w), c,$

$$\begin{aligned} &V(g'_c(\cdot)|_{p_1}) \leq U(g'_c(\cdot)|_{p_1}) \implies \\ &\sum_{a \in A} V(g'_a(\cdot)|_{p_1}) \cdot p_1 g \leq \sum_{a \in A} U(g'_a(\cdot)|_{p_1}) \cdot p_1 g \implies \\ &\sum_{w \in W} \left(\sum_{a \in A} u(c, l) p_1 \right) g(w) + \beta \sum_{a \in A} V(g'_a(\cdot)|_{p_1}) \cdot p_1 g \\ &\leq \sum_{w \in W} \left(\sum_{a \in A} u(c, l) p_1 \right) \implies \\ &H^{p_1}V(g) \leq H^{p_1}U(g) \implies \\ &H^{p_1}V(g) \leq H^{p_2}U(g) \implies \\ &HV(g) \leq HU(g) \implies \\ &HV \leq HU. \end{aligned}$$

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 (5)-(8), 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 α – vectors, $\alpha : W \rightarrow R$, 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 (5)-(8). For planning horizon $n = 0$, I have only to take into account the immediate expected rewards and thus I have that:

$$V_0(g) = \max_{p \in \Gamma} \left[\sum_{w \in W} \left(\sum_{a \in A} u(c, l) p \right) g(w) \right] \quad (14)$$

and therefore if I define the vectors

$$\{\alpha_0^i(w)\}_i \equiv \left(\sum_{a \in A} u(c, l) p \right)_{p \in \Gamma} \quad (15)$$

I have the desired

$$V_0(g) = \max_{\{\alpha_0^i(w)\}_i} \langle \alpha_0^i, g \rangle \quad (16)$$

where $\langle \cdot, \cdot \rangle$ 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 (??):

$$V_n(g) = \max_{p \in \Gamma} \left[\sum_{w \in W} \left(\sum_{a \in A} u(c, l) p(c, l|w) \right) g(w) + \beta \sum_{w \in W} \sum_{a \in A} (V_{n-1}(g'_a(\cdot)|_a)) p(c, l|w) g(w) \right] \quad (17)$$

by the induction hypothesis

$$V_{n-1}(g(\cdot)|_a) = \max_{\{\alpha_{n-1}^i\}_i} \langle \alpha_{n-1}^i, g'_a(\cdot) \rangle \quad (18)$$

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

$$V_{n-1}(g'_a(\cdot)) = \max_{\{\alpha_{n-1}^i\}_i} \sum_{w' \in W} \alpha_{n-1}^i(w') \left(\sum_{w \in W} \sum_{a \in A} T(\cdot; w, c, l) \frac{\Pr(w, c, l)}{\Pr(c, l)} \right) \quad (19)$$

With the above:

$$\begin{aligned}
V_n(g) &= \max_{p \in \Gamma} \left[\begin{aligned} &\sum_{w \in W} \left(\sum_{a \in A} u(c, l) p \right) g(w) + \\ &+ \beta \max_{\{\alpha_{n-1}^i\}_i} \sum_{w' \in W} \alpha_{n-1}^i(w') \left(\sum_{w \in W} \left(\sum_{a \in A} \frac{T(\cdot; w, c, l)}{\Pr(c, l)} \cdot p \right) g(w) \right) \end{aligned} \right] \\
&= \max_{p \in \Gamma} \left[\langle u(c, l) \cdot p, g(w) \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c, l)} \max_{\{\alpha_{n-1}^i\}_i} \left\langle \sum_{w' \in W} \alpha_{n-1}^i(w') T(\cdot; w, c, l) \cdot p, g \right\rangle \right]
\end{aligned} \tag{20}$$

At this point, it is possible to define

$$\alpha_{p,a}^j(w) = \sum_{w' \in W} \alpha_{n-1}^j(w') T(\cdot; w, c, l) \cdot p. \tag{21}$$

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

$$V_n(g) = \max_{p \in \Gamma} \left[\langle u(c, l) \cdot p, g \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c, l)} \max_{\{\alpha_{p,a}^j\}_j} \langle \alpha_{p,a}^j, g \rangle \right], \tag{22}$$

and define:

$$\alpha_{p,a,g} = \arg \max_{\{\alpha_{p,a}^j\}_j} \langle \alpha_{p,a}^j, g \rangle. \tag{23}$$

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

$$\begin{aligned}
V_n(g) &= \max_{p \in \Gamma} \left[\langle u(c, l) \cdot p, g \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c, l)} \langle \alpha_{p,a,g}, g \rangle \right] \\
&= \max_{p \in \Gamma} \left\langle u(c, l) \cdot p + \beta \sum_{a \in A} \frac{1}{\Pr(c, l)} \alpha_{p,a,g}, g \right\rangle.
\end{aligned} \tag{24}$$

Now

$$\{\alpha_n^i\}_i = \bigcup_{\forall g} \left\{ u(c, l) \cdot p + \beta \sum_{a \in A} \frac{1}{\Pr(c, l)} \alpha_{p,a,g} \right\}_{p \in \Gamma} \tag{25}$$

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 (15), the set of vectors $\{\alpha_0^i\}_i$,

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

is discrete. For the general case, observe that for discrete controls and assuming $M = |\{\alpha_{n-1}^j\}|$, the sets $\{\alpha_{p,c}^j\}$ 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

$\theta = 2, \{\gamma = 1, \eta = 0\}$

	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	3.55	1.29	0.92
Labor (l)	2.75	1.12	0.88
Wealth (w)	5.73	2.11	0.94
Information Flow (κ)	0.53	0.91	

Cross-Correlation

	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.92	1	0.86	0.89	0.75	0.47
(l)	0.88	0.89	0.88	1	0.72	0.41
(w)	0.30	0.47	0.37	0.41	0.94	1

Table 10a: Statistical properties of the Model, $M(2,1,0)$

$\theta = 0.2, \{\gamma = 1, \eta = 0\}$

	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	4.05	1.11	0.95
Labor (l)	2.98	0.97	0.94
Wealth (w)	5.84	1.86	0.93
Information Flow (κ)	0.98	0.63	

Cross-Correlation

	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.95	1	0.90	0.90	0.77	0.84
(l)	0.87	0.90	0.94	1	0.82	0.79
(w)	0.87	0.84	0.75	0.79	0.93	1

Table 10b: Statistical properties of the Model, $M(0.2,1,0)$

$$\theta = 0.02, \{\gamma = 1, \eta = 0\}$$

	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	4.26	0.75	0.97
Labor (l)	3.76	0.73	0.77
Wealth (w)	6.06	1.21	0.89
Information Flow (κ)	1.52	0.34	

Cross-Correlation						
	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.97	1	0.86	0.94	0.48	0.88
(l)	0.86	0.94	0.77	1	0.76	0.87
(w)	0.88	0.78	0.77	0.87	0.89	1

Table 10c: Statistical properties of the Model, $M(0.02,1,0)$

$$\theta = 0.2, \{\gamma = 3, \eta = 0\}$$

	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	4.11	1.31	0.93
Labor (l)	3.90	0.92	0.91
Wealth (w)	5.88	1.84	0.90
Information Flow (κ)	1.10	0.45	

Cross-Correlation						
	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.93	1	0.90	0.85	0.64	0.72
(l)	0.76	0.85	0.91	1	0.66	0.75
(w)	0.50	0.62	0.58	0.55	0.90	1

Table 10d: Statistical properties of the Model $M(0.2,3,0)$

$$\theta = 0.2, \{\gamma = 1, \eta = 1\}$$

	Mean	St.Dev	$\rho(x_t, x_{t-1})$
Consumption (c)	4.09	0.99	0.94
Labor (l)	2.90	0.74	0.95
Wealth (w)	5.07	1.31	0.97
Information Flow (κ)	1.01	0.59	

Cross-Correlation						
	$c(-1)$	c	$l(-1)$	l	$w(-1)$	w
(c)	0.94	1	0.88	0.90	0.73	0.67
(l)	0.95	0.90	0.95	1	0.44	0.54
(w)	0.56	0.67	0.64	0.64	0.97	1

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

8.2 Figures

Comparing consumption for different utilities, $\theta=0.2$

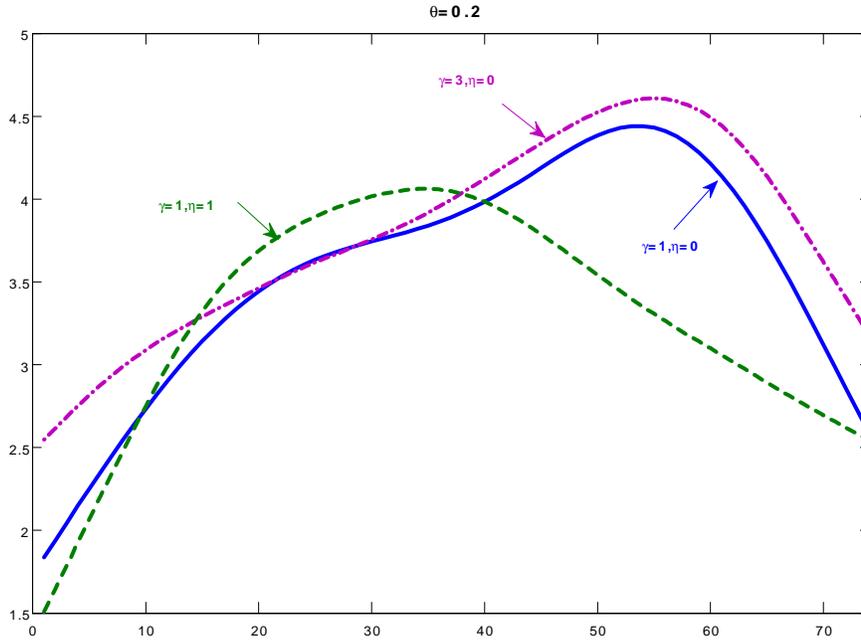


Figure 7 :Blue: $\gamma=1, \eta=0$; Green: $\gamma=1, \eta=1$; Violet: $\gamma=3, \eta=0$.

Comparing labor and wealth for different utilities, $\theta=0.2$

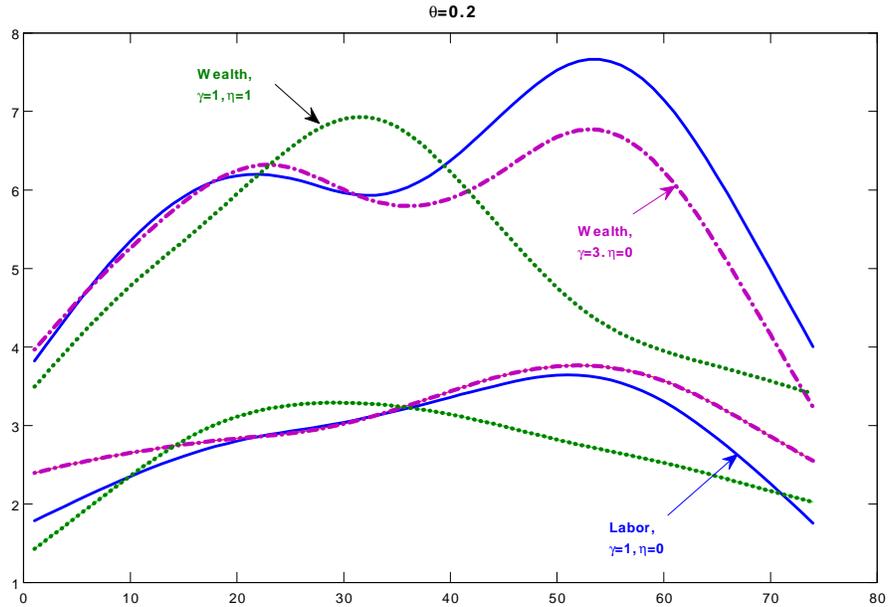


Figure 8 :Blue: $\gamma=1, \eta=0$; Green: $\gamma=1, \eta=1$; Violet: $\gamma=3, \eta=0$.

Comparing savings for different utilities, $\theta=0.2$

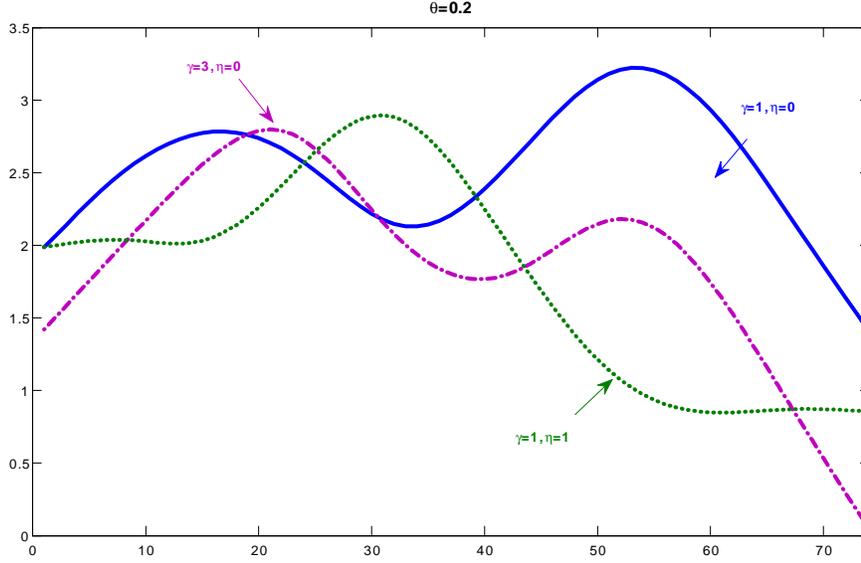


Figure 9 :Blue: $\gamma=1, \eta=0$; Green: $\gamma=1, \eta=1$; Violet: $\gamma=3, \eta=0$.

Comparing wealth, consumption, savings and income for different θ 's, $u(c,l)=\log(c)-\alpha l$

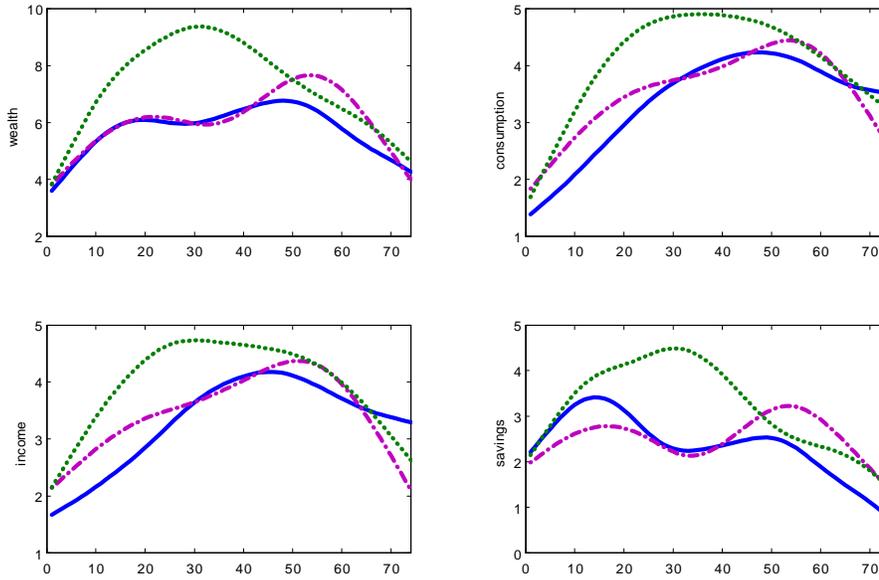


Figure 10 :Blue: $\theta=2$; Violet: $\theta=0.2$; Green: $\theta=0.02$.

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 $\Omega_c \equiv (c_{low} \ c_{high})$. 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:

$$c \leq \bar{w} + s\mathbf{1}_{\{l=1\}} \quad (26)$$

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:

$$\alpha c \leq s \quad (27)$$

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 (26), i.e., $\{(c_{high}, l = 1), (c_{low}, l = 1), (c_{low}, l = 0)\}$. 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

$C, L \setminus S$	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.

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

The first order condition are

$\partial z_1 :$

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

$\partial z_2 :$

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

$\partial z_3 :$

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

The set of first order conditions yield a system of simultaneous transcendental 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 $\omega_i \equiv u(c_i, l = 1) - u(c_{low}, l = 0)$ where $i \equiv (c_{high}, c_{low})$. Then the solution for z_i is given by

$$z_i = \frac{-\lambda / (\omega_i + p)}{W((\lambda / (\omega_i + p)) e^{-\lambda(1+p)/\omega})}. \quad (28)$$

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

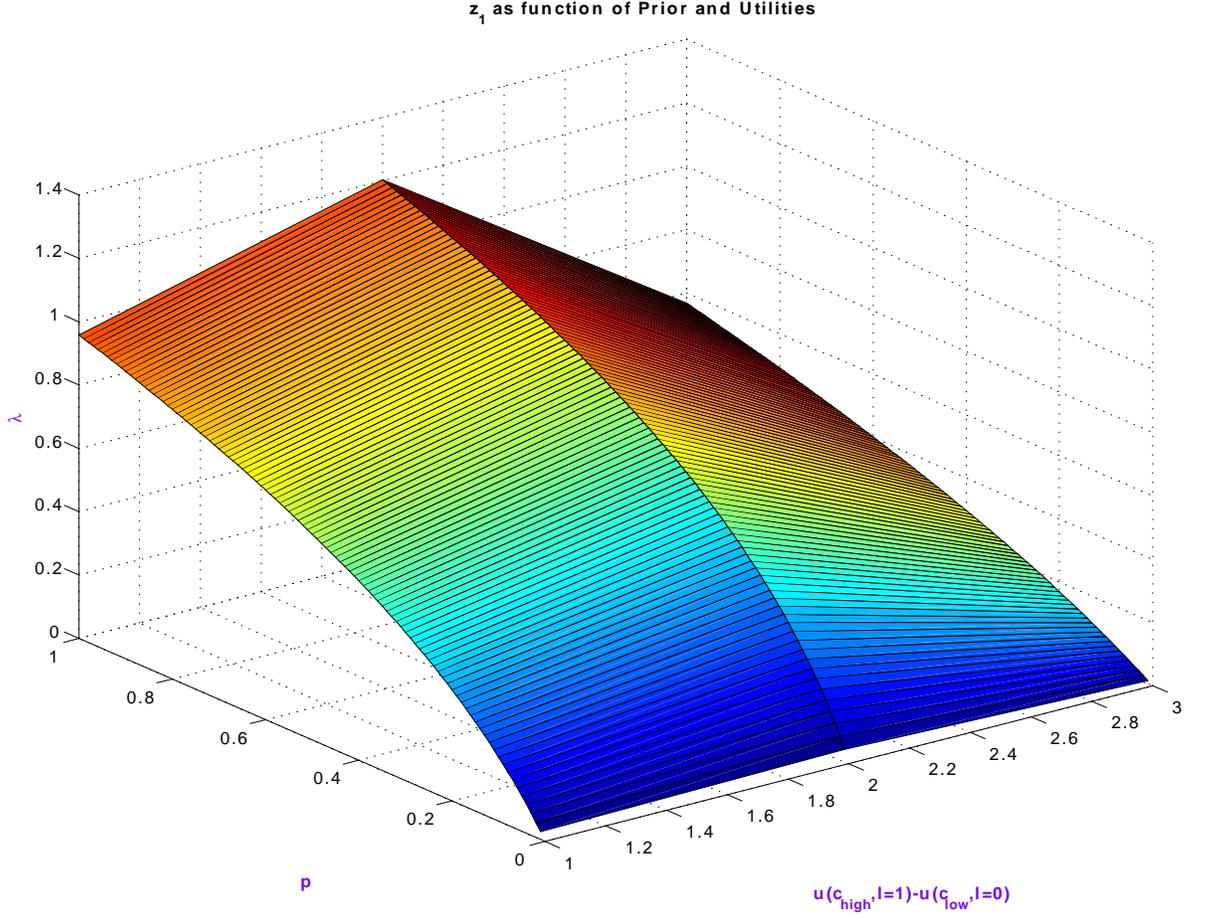


Figure 11 :An analytical solution

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 $V(g'_{a_j}(\cdot)) = 0$.
- 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) .

→ 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.