# Rationally Inattentive Macroeconomic Wedges

## Antonella Tutino

## Research Department
## Working Paper 1005

**Federal Reserve Bank of Dallas**

# Rationally inattentive macroeconomic wedges

Antonella Tutino*

Federal Reserve Bank of Dallas

September 8, 2010

**Abstract**

This paper argues that the solution to a dynamic optimization problem of consumption and labor under finite information-processing capacity can simultaneously explain the intertemporal and intratemporal labor wedges. It presents a partial equilibrium model where a representative risk adverse consumer chooses information about wealth with limited attention. The paper compares ex-post realizations of models with finite and infinite capacity. The model produces macroeconomic wedges and measures of elasticity consistent with the literature. These findings suggest that a consumption-labor model with information-processing constraints can explain the difference between predicted and observed consumption and employment behavior.

JEL: C44; C63; D81; D91.

Keywords: *Finite Shannon capacity, macroeconomic wedges, savings decisio*n.

# Introduction

Rational expectation models of consumption-leisure choice predict that people maximize their utility by balancing contemporaneous and future consumption and leisure. The trouble is that U.S. data on consumption and employment do not agree with the optimality conditions that this theory generates. Mankiw, Rotemberg and Summers (1985) were the first to formally test these optimality conditions against U.S. data. They concluded that the realizations from a stochastic dynamic optimization problem of a rational representative agent do not match the observed consumption and work behavior in the U.S. The macroeconomic literature[1] has analyzed extensively the ways in which the solution of a rational expectation consumption-labor problem departs from the data in search of regularities and explanations. Despite little consensus on the causes, the empirical and the theoretical literature seem to agree that the differences between data and theory occur on both the intratemporal and inter-temporal margins of optimization.

The starting point of the paper is the idea that rational inattention theory might offer a way to reconcile these differences by recognizing that people allocate little attention to changes in wealth at high frequencies and, as a result, change their intertemporal consumption and leisure plans less frequently than rational expectation theory predicts. Moreover, rational inattention predicts that people react to changes in wealth over time by varying their plans. This way of changing consumption and labor behaviors needs not to be the same as the long-run behavior postulated by the intratemporal first order condition of a rational expectation model of optimal behavior.

The purpose of the paper is to show that the ex-post realizations from a stochastic dynamic optimization problem of a rational inattentive (Sims, 2003, 2006) representative agent can contemporaneously account for the intertemporal labor and consumption wedges together with the intratemporal labor wedge documented in the literature.

---

[1]See, e.g., Mankiw, Rotemberg and Summers (1985), Barro and King (1984), Hall (1997), Chari, Kehoe and McGrattan (2007), and Chang and Kim (2009). For a survey, see Shimer (2009).

The paper introduces Shannon's information-processing constraints to a model of consumption-labor choice where the only source of exogenous stochasticity comes from wages. Given his uncertainty on how changes in wages impact wealth, the consumer selects information on wealth functional to his consumption and labor decisions through a finite capacity Shannon's channel. Consistent with fully dynamic rational inattention problems (Sims 2006, Tutino 2009), the paper assumes general ex-ante uncertainty and utility specifications. Discretization and dynamic programming solve the model producing consumer's optimal choice of information which, in turn, defines consumption and labor plans. Simulating the model, the paper shows that these plans are, at least qualitatively, consistent with observed behavior of consumption and labor in the U.S. for both the long-run and the short-run. In particular, when applied to the three margins of optimizations postulated by a rational expectation model, the time series of consumption and labor generate differences between theory and data similar to the intratemporal and intertemporal wedges documented by the literature.

The paper takes the optimality conditions from the rational expectation model and defines these wedge as the difference between the realizations of consumption and labor that comes from the model and those generated from a rational inattention model. The analysis of these differences proceeds in three ways. The first way is to compute the wedges from the definition by feeding to the optimality conditions the time series of consumption and labor generated by the rational expectation model and rational inattention models. This step gives a statistical assessment of the size and scope of the intertemporal and intratemporal wedges as a function of information-processing constraints. The second way is to study the differential responses of consumption and labor to permanent and transitory shocks to wages. The rational inattention model-generated impulse-response functions show that a temporary expected positive shock to wage produces large variation in the short-run labor supply. The same model shocked with a permanent expected positive innovation to wage produces significant variation in the long-run consumption and minimal variation in the long-run labor supply. These results appear robust to different specifications of utility and

2

information-processing constraints. Finally, the paper measures the model-generated short-run and long-run elasticities of labor supply. In particular, the paper asks what estimates of short-run and long-run elasticities an econometrician would produce if he analyzes data coming from a rational inattention representative agent economy through the lens of the differential rational expectation optimality conditions. The results from this exercise are (1) a backward bending labor supply curve and (2) an extremely elastic short-run labor supply curve. These estimates are similar to those first documented by Mankiw, et. al. (1985).[2]

The outline of the paper proceeds as follows. Section 2 presents the model under infinite Shannon's information-processing constraints –i.e., rational expectation– and finite information-processing constraints –i.e., rational inattention– and defines the macroeconomic wedges between the two specifications. Section 3 derives the model's predictions. Section 4 computes wedges and elasticities and Section 5 concludes. Appendix A provides details on the statistical methodology. Appendices B and C are in the journal's archive. Appendix B collects additional figures and statistics and Appendix C presents a pseudocode.

# 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. Finally, I characterize the a-temporal and intertemporal wedges that derive from the comparison of the two models.

## A version of the model under infinite processing capacity

The economy is populated by a representative household who maximizes the expected discounted value of his utility. Utility, $u(C, L)$, is defined over a consumption good, $C$, and

---

[2]Mankiw et. al. (1985) report similar findings on estimates of elasticities of labor supply fitting U.S. data to a stochastic dynamic rational expectation model of consumption-labor choice.

labor, $L$ and it is strictly increasing and concave in $C$ and decreasing and concave in $L$, $\lim_{C \to 0} u_C(C, L) = +\infty, \forall L \in [0, 1]$, $\lim_{L \to 0} u_L(C, L) = -\infty \; \forall C \geq 0$. Moreover, I assume that utility is homogeneous and additively separable. In particular:

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

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

Each period, the consumer faces a stochastic real wage, $s$, in exchange for his labor effort. Wage follows a stationary i.i.d. Markov process. [3] Consumer's wealth, $W$, evolves according to previous' period savings $(W - C)$, augmented by 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 i.i.d. processes, 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

---

[3] Given that the point of the paper is to characterize the macroeconomic wedges between a rational expectation model and a model with rational inattention in the simplest possible framework, I assume that wages are i.i.d..

with respect to consumption and labor imply the following contemporaneous relation

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

So long as both $\eta$ and $\alpha$ are finite and with a desire to smooth consumption implied by (1), condition (S) implies that labor will change through time reflecting changes in wages. The intertemporal optimal condition for consumption is:

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

and the equivalent for labor:

$$1 = -E_t \left[ \left( \frac{L_{t+1}}{L_t} \right)^{\eta} \cdot \frac{s_t}{s_{t+1}} \right]. \tag{EL}$$

To match the joint behavior of per capita consumption and per capita hours that we observe in U.S. data, the model should produce (1) cyclical movements: procyclical behavior of per capita consumption and per capita hours worked–i.e., equation (S) calls for a high intertemporal elasticity of labor supply –; (2) secular movements: labor supply's response to permanent changes in wages is negligible while consumption's responses are significant and essentially match the changes in wages–i.e., equation (EL) calls for a low intertemporal elasticity of labor supply–.

As Mankiw, et. al. (1985) noted, the first order conditions (S)-(EL) cannot account for facts (1) and (2) simultaneously. They fit the optimality conditions (S)-(EL) to U.S. data and find that U.S. data do not support the model. This negative result prompts the authors to question the validity of stochastic dynamic optimization as useful framework to match data on consumption and labor.

# A version of the model under finite processing capacity

Under *Rational Inattention Theory* (Sims, 1998, 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. The assumption that people process information at finite rate implies a profound modification of the problem (2)-(5). First, with finite-processing capacity wages and, in turn, wealth cannot be observed perfectly at any point in time. As a result, the consumer starts his optimization problem with an idea of what his wealth is. I model formally this concept by characterizing this "idea" as a set of possible events concerning wealth $W$ with associated probability distribution, $b(W)$. This distribution constitutes the new state of the problem. Second, given that the knowledge of wealth is stochastic, before processing any information, the decisions $A = C, L$ of consumption, $C$, and hours worked, $L$, ought to be random variables. Third, given the stochastic knowledge of decisions and state and limited-processing capacity, to solve the optimization problem the consumer needs a strategy that relates information about wealth to his decisions. In other words, he needs to choose signals about wealth functional to sharpen his consumption and labor plans. I model formally this concept by having the consumer choose the joint distribution of labor, consumption and wealth, $p(\cdot_w; \cdot_a)$, as control variable for the state $b(W)$. 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 information content of the signal cannot exceed consumer's processing capacity. Forth, the consumer evaluates the outcome of his strategy by observing his consumption profile $(c^*)$, and labor supply $(l^*)$. I model this by assuming that the consumer draws from the optimal distribution $p^*(\cdot_w; \cdot_a)$ a realization $(a = [c, l])$.
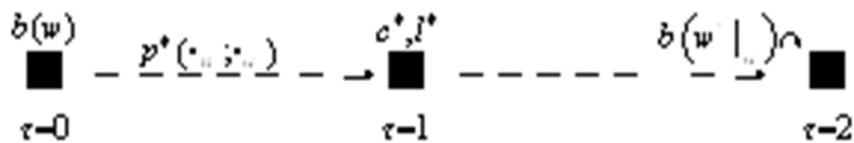


**Figure 1**:A typical day of a rationally inattentive person.

Fifth, the consumer observes the outcomes of his choices and uses the observation to update rationally his knowledge of wealth. Formally, given an outcome $a$ and using Bayes' rule, consumers' knowledge of next period's wealth is embedded in the posterior $b\left(w'|_a\right)$. Figure 1 describes the events.

## Statement and Recursive Formulation of Consumer's 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 a finite-capacity consumer. Second, I turn to the information-flow constraint, key of the model. Third, I present the objective function and cast the problem into a recursive formulation. The structure of the economy follows closely the one of Tutino (2009) to which I refer for the mathematical details. For completeness, Appendix A rigorously proves that the problem admits a recursive formulation and that the resulting Bellman equation is a contraction.

### Budget Constraint and Update

Recall that $C$ denotes consumption of the good in the economy and $L$ is the amount of labor supplied. I collect the actions at time $t$ in the set $A_t \equiv \{C_s, L_s\}_{t \leq s \leq \infty}$. The consumer is limited in his choices by the same budget constraint as in (3) and reported here for convenience:

$$W_{t+1} = R_t \left(W_t - C_t\right) + s_t L_t \tag{6}$$

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. process. Wages are the only exogenous source of stochasticity in the model. The consumer wishes to reduce uncertainty about the linear combination of savings and labor income as displayed in (6). Since information is 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 the consumer makes consumption and labor decisions knowing that the mean of the wages is fixed at $\bar{s}$. This knowledge is embedded in a prior over the possible realizations of wealth, $g(w_t)$. The consumer updates rationally his belief through signals on wealth and observation of past consumption and work behaviors.

Let $a \equiv \{c, l\}$ be a particular behavior of consumer where $c$ is a specific realization of the random variable $C$ and, similarly, $l$ is a specific outcome of the random variable $L$. The posterior of wealth $(w')$ conditional on observing a $a = \tilde{a}$ follows by Bayes' law:

$$g'\left(w'|_{\tilde{a}}\right) = \int_w T\left(w'; w, \tilde{a}\right) \cdot p\left(w|\tilde{a}\right) dw \tag{7}$$

where $T(w'; ., .)$ derives from the dynamic of wealth (6).The operator $T(w'; w, \tilde{a})$ assigns probabilities to $w'$ conditional on the value of $\tilde{a}$ and $w$. Since values of $w$ are not observable, the operator $T(.)$ 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:

$$T\left(w'; w, \tilde{a}\right) \equiv E\left(W'\right) = R \int_w \left(w - \tilde{c}\right) dw + \bar{s}\tilde{l}. \tag{8}$$

The distribution $p\left(w|\tilde{a}\right)$ takes into account the potential noise in the current observation of the state that arises from choosing the signal $p(w; \tilde{a})$. This noisy observation is carried over one period ahead to infer next period's state.

**Choice Variable**

As stated in Section 2.2, before processing any information about wealth $(W)$, consumption $(C)$ and labor $(L)$ are random variables from the consumer's perspective. Thus, the consumer cannot solve his optimization problem without relating wealth to his behavior $(C$ and $L)$. Since the consumer starts his life with a probabilistic knowledge of $W$, mapping $W$ into $C$ and $L$ translates into finding a joint relation among wealth, consumption and labor

that matches information about wealth to consumption and labor. The selection of this information, that is, the joint probability distribution of wealth, consumption and labor, $p\left(\cdot_w, \cdot_c, \cdot_l\right)$, is key in the optimization of the consumer since it affects current beliefs and updates.

To clarify this point, suppose that the consumer can process information at infinite rate. In this case, the optimal $p\left(\cdot_w, \cdot_c, \cdot_l\right)$ will be degenerate assigning to each value of $w$ one value for $c^*\left(w\right)$ and $l^*\left(w\right)$, as in the solution (S)-(EL). By contrast, suppose that the consumer has extremely limited information-processing capacity. In this case, he might choose to use the capacity to assess the limits of his wealth and to keep consumption and labor fairly constant and at or below those limits.[4] When the information-processing effort lies in between this two extreme cases, optimizing consumers set $p\left(\cdot_w, \cdot_c, \cdot_l\right)$ such that the conditional probability of wealth given consumption and labor -i.e., the posterior of wealth- is as close to wealth as possible given the information constraint and the preference spelled out in the utility function.[5]

**Information Constraint**

For people with limited information-processing capacity, attention is a scarce resource. To model the technology that makes this resource limited I use Shannon's mutual information[6] between the random variables $W$ and $A$. Mutual information defines the capacity of the channel and depends only on the joint distribution of $W$ and $A$ for a given belief $g\left(W\right)$. In the context of the model, Shannon's capacity captures the ability of consumers to interpret news about their wealth, thereby regulating the speed of reaction of their behavior to this news. Formally, Shannon's capacity measures the maximum reduction in uncertainty as the

---

[4]The behavior described is certainly true for a risk-averse consumer with very limited capacity as illustrated in section 4. I have not investigated the cases of risk-lover or risk-neutral consumers.

[5]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 Sims (2006) and Tutino (2010) for a discussion of advantages and disavantages of moving into a fully dynamic rational inattention model with respect to the LQG framework.

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

difference between the initial uncertainty -entropy of $W$- and the knowledge of the variable $W$ provided by the observation of $A$ -i.e., conditional entropy of $W$ given $A$-. [7]

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 $A$-.[8] In the model, such a cost is denoted by $\theta$.[9] 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 :

$$\kappa_t = \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right) = \int p\left(w_t, a_t\right) \log\left(\frac{p\left(w_t, a_t\right)}{\left(\int p\left(\tilde{w}_t, a_t\right) d\tilde{w}_t\right) g\left(w_t\right)}\right) dw_t da_t \qquad (9)$$

The expression in (9) says that the uncertainty that the consumer can reduce about wealth through observation of consumption and labor supply is at most $\kappa$ 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 very informative about wealth. This strategy results in a policy function for consumption, labor and wealth that depends on the -now observable- value of wealth. On the other extreme, with almost no processing capacity, a consumer might want to assign all the probability to a particular value of $A$. This choice makes the variables $W$ and $A$ independent of each other, $\left(\mathcal{I}\left(p\left(\cdot_w; \cdot_a\right)\right) \to 0\right)$. In the intermediate case, the person 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.

---

[7] Since mutual information is a function of the joint distribution of $W$ and $A$ for a given belief, it is applicable regardless of the nature and characteristics of the channel.

[8] The assumption of having a shadow cost on information-processing constraint reduces the computational burden. Moreover, note that having a shadow cost associated to information-processing constraint and $\kappa_t = \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right)$ holding with equality, is isomorphic to assuming a bound on the maximum amount of capacity while having the constraint holding with inequality, that is $\bar{\kappa} \geq \mathcal{I}_t\left(p\left(\cdot_w, \cdot_a\right)\right)$, $\forall t$. The isomorphism comes from realizing that there exists a one-to-one mapping between $\bar{\kappa}$ and the shadow cost of the channel, $\theta$. 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 (2010).

[9] This way of modeling information-processing constraints is consistent with Sims (2006, 2009) and Macrowiak and Wiederholt (2009).

## Objective

Consumer's problem is to maximize over an infinite planning horizon expected utility of consumption and leisure discounted at factor $\beta < 1$. Let $\theta$ be the (fixed and exogenous) shadow cost of processing information $\kappa$ in (9). Moreover, let utility be:

$$u(c,l) = \log c - \frac{\alpha}{\eta+1} l^{\eta+1} \tag{10}$$

where $\eta$ is the inverse of Frisch elasticity of labor supply - defined as $\varepsilon$- and $\alpha \in [0,1]$ is a constant disutility associated to labor. The control for agent's maximization is the joint distribution of actions and wealth, $p(w,a)$, that solves : [10]

$$V(g(w)) = \max_{p(w,a) \in \mathcal{D}} \sum_w \sum_a u(c,l) p(w,a) - \theta\kappa + \beta \left( \sum_w \sum_a (V(g'(w'|_a))) p(w,a) \right) \tag{11}$$

subject to:

$$\kappa = \mathcal{I}(p(\cdot_w, \cdot_a)) = \sum_w \sum_a p(w,a) \log \left( \frac{p(w,a)}{(\sum_{\tilde{w}} p(\tilde{w},a)) g(w)} \right) \tag{12}$$

together with the rational update (7) and the requirement that $p(w,a) \in \mathcal{D}$, where $\mathcal{D}$ is the set of all distributions that satisfy:

$$\sum_a p(\tilde{w},a) = g(\tilde{w}) \tag{13}$$

$$0 \le p(\cdot_w; \cdot_a) \le 1, \; \forall (w,a) \tag{14}$$

$$\sum_w \sum_a p(w,a) = 1. \tag{15}$$

---

[10] For a formal prove that the infinite problem of the household can be cast into a recursive Bellman program see Appendix A.

In addition (3) and (5) are imposed. Taking first order condition with respect to $p(w, a)$ results in [11]

$\partial p(w, a):$

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

$$+\beta \left( \sum_w \sum_a \left[ \frac{\partial \left( V \left( g'\left( w'|_{a_t} \right) \right) \right)}{\partial g\left( w'|_{a_t} \right)} \frac{\partial g\left( w'|_a \right)}{\partial p(w, a)} \right] p(w, a) \right) \quad (16)$$

where

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

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

The expression in (16) reveals that in selecting the optimal $p(w, a)$ the consumer evaluates the impact that his choice has on the current level of his value function, $u(c, l) + \beta V(\bullet|_a)$, as well as how his choice shapes up the perception of future values and beliefs, $\left[ \partial \left( V \left( g'\left( w'|_{a_t} \right) \right) \right) / \partial g\left( w'|_{a_t} \right) \right] \times \left[ \partial g\left( w'|_a \right) / \partial p(w, a) \right]$. To my knowledge, expressions like (16) do not admit an analytical solution except for some particular cases. [12] Define the solution to the optimization problem of the consumers as the distribution $p^*\left( \cdot_w; \cdot_a \right)$. The realized outcomes $a_t = \left( \{c_t\}, \{l_t\} \right)$ are then drawn from the optimal joint $p^*\left( \cdot_w; \cdot_c, \cdot_l \right)$.

---

[11] Note that the first order condition in (16) is valid for $\theta > 0$. If $\theta = 0$, then the probabilities $g(w)$ and $p(\cdot_w, \cdot_a)$ are degenerate. In this case, $\mathcal{I}(p(\cdot_w; \cdot_a)) = 0$, and using Fano's inequality (Thomas and Cover 1991), $c(\mathcal{I}(p(\cdot_w; \cdot_a))) = c(w)$ and $l(\mathcal{I}(p(\cdot_w; \cdot_a))) = l(w)$. This result makes the first order conditions for this case equal to the full information solution.

[12] The solution of a static problem with log utility corresponds to a Lambert W function which in general does not have a close form expression. Analytical expressions are possible under some restriction on the utility function and the ex-ante shape of uncertainty. One such case is within the Linear Quadratic Gaussian framework. For others, see Matejka and Sims (2009).

## Wedges

When the consumer chooses the joint distribution $p^* \left( \cdot_w; \cdot_c, \cdot_l \right)$, he is ultimately choosing a metric under which he forms expectations on future values of wealth. To be more specific, under infinite-processing capacity, the information set upon which expectations are conditioned corresponds to the set $\Omega_t$ that includes current and past realization of wealth, consumption and labor. By contrast, with finite-processing capacity, expectations are conditioned on a set $\mathcal{Z}_t$ that includes current and past beliefs about wealth, current and past optimal choices, $p \left( \cdot_w; \cdot_a \right)$, and realizations of consumption and labor. If the channel transmits at finite rate there will be a difference between the optimality conditions (S)-(EL) evaluated under the expectations $E_t \left( . | \Omega_t \right) \equiv E_t \left( . \right)$ and $E_t \left( . | \mathcal{Z}_t \right) \equiv E_{p^*} \left( . \right)$.

Consider the static first order condition (S) and let $m \left( a, w \left( s \right) \right) \equiv \left( -\partial U \left( a, w \right) / \partial L / \partial U \left( a, w \right) / \partial C \right)$ be the marginal rate of substitution. Define $\Lambda^s$ as the static wedge that makes the first order condition evaluated at an optimal strategy $p^*(w, a)$, that is:

$$1 = \left( 1 - \Lambda^s \right) E_{p^*} \left[ \frac{1}{s} * m \left( a, w \left( s \right) \right) \right],$$

then,

$$\Lambda^s \equiv 1 - \left( E_{p^*} \left[ \frac{1}{s} * m \left( a, w \left( s \right) \right) \right] \right)^{-1}. \tag{17}$$

Similarly, let $\gamma_t^C \left( a, w \left( s \right) \right) \equiv \left( \frac{\partial U(a,w)/\partial C_{t+1}}{\partial U(a,w)/\partial C_t} \right)$, and, using (EC) the consumption intertemporal wedge is:

$$\Lambda_t^{EC} \equiv 1 - \left[ E_{p^*} \left( \gamma_t^C \left( a, w \left( s \right) \right) \right) \right]^{-1} \tag{18}$$

and finally, from (EL) the labor intertemporal wedge is:

$$\Lambda_t^{EL} \equiv 1 + \left[ E_{p^*} \left( \gamma_t^L \left( a, w \left( s \right) \right) \right) \right]^{-1} \tag{19}$$

where $\gamma_t^L \left( a, w \left( s \right) \right) \equiv \left( -\frac{\partial U(a,w)/\partial L_{t+1}}{\partial U(a,w)/\partial L_t} \frac{s_t}{s_{t+1}} \right).$

As noted, in general there is no close form solution for $p^*(.;.)$, which makes the computation of expectations $E_{p^*(c,w)}$ and, in turn, an analytical characterization of the wedges (17)-(19) very difficult. [13] However, as pointed out by Altonji (1982), a viable alternative to knowing people's expectations is to look at behavior which, together with the theory, reveals what people were expecting. This paper follows Altonji's intuition by evaluating the expectations derived from the solution of the rational inattention model (11) - (14) through the simulated time series of consumption, labor and wealth. This way of computing expectations serves two purposes. The first is to derive the intertemporal and intratemporal wedges (17)-(19) by comparing the behavior under rational inattention and rational expectations. The second purpose is to evaluate short-run and long-run elasticities of intertemporal substitution feeding the model-generated time series into the totally differentiated expressions (S)-(EL).[14]

Before turning to the study of the wedges and the elasticities, next section solves the model and illustrates some properties of the solution. A reader not interested in technical aspects and robustness results may want to skip directly to Section 4.

# Solution and predictions of the model

## Optimal conditional distribution

Figures 2a and 2b show the optimal conditional distribution of consumption and labor[15] respectively for a given value of wealth - $w = 1$, top panel $w = 4$, medium panel and $w = 8$,

---

[13] As an example of the complexity involved, to evaluate the expectations in (17), one needs to compute:
$$\int s_t m(a,w(s))g(w(s))d\mu(w(s)) - \int s_t m(a,w(s))\delta(a-\tilde{a}^*(w))d\mu(w(s))$$
where $\tilde{a}^*(w)$ is the combination of $(L,C)$ that satisfy condition (S), $\delta(.)$ is the Dirac function equal to 1 when $a = \tilde{a}$ and zero elsewhere, and $\mu(w(s))$ is the measure under which we evaluate the integral that allows for discreteness in the distribution.

[14] This way of computing elasticities is used by Mankiw, et. al. (1985). A similar venue is taken by Chang and Kim (2009).

[15] The optimal conditional distributions in Figures 2a and 2b are calculated as follows. For a given value of wealth, $w = w^*$, the optimal conditional distribution of consumption is $p(c|w^*) = \sum_{l \in L} p^*(c,l|w^*)$, with $p^*(c,l|w^*)$ coming from the solution of the optimization problem (11)-(5) and $L$ being the set of possible values of labor, $l$. A similar expression holds for $p(l|w^*)$.

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.[16] Consider first the optimal conditional distribution of consumption when wealth is low -Figure 2a, top panel-. For $w = 1$, the optimal signal for a person with $\theta = 0.2$ - blue bars in the picture- places high probability mass on low values of consumption $(c = 0.7)$ but it also allows for the possibility of higher consumption - $c = 2.3$ and $c = 3.1$- sustained by labor income. As the top panel of Figure 2b 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 2a and 2b- 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 worked 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 focused on the medium values of the support with a peak at $l = 2.8$ with

---

[16]In Appendix B, the values reported for the statistics of consumption, labor, wealth and information flow are average across simplex-points and Monte Carlo runs.
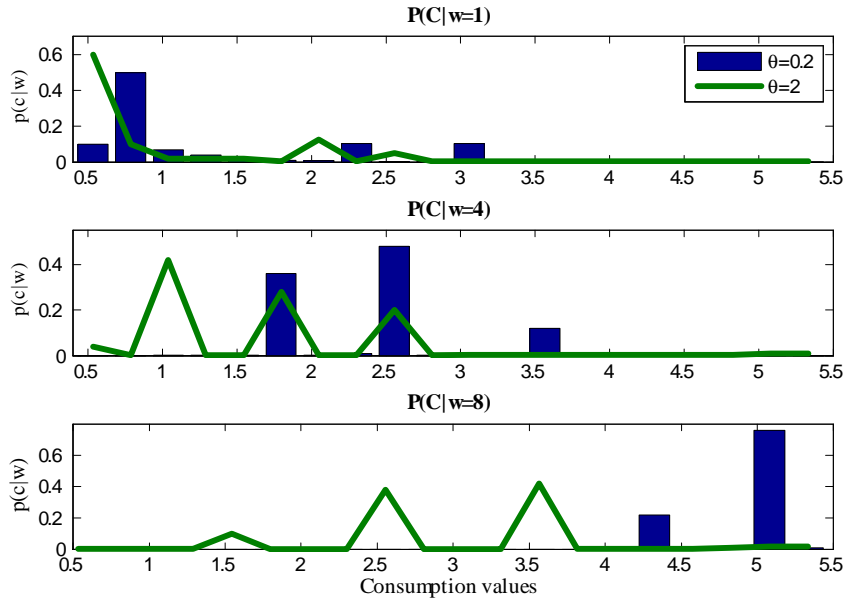
probability 0.42.



**Figure 2a:** Optimal conditional distribution of consumption for $\theta$=0.2 (bar) and $\theta$=2 (line)

Now consider an agent with the same preferences as the previous one but $\theta = 2$. Figure 2a and 2b show that a $\theta = 2$-type has more spread distributions of consumption and labor than a $\theta = 0.02$-type. The noisier behaviors of labor and consumption are due to the fact that people with $\theta = 2$ have 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
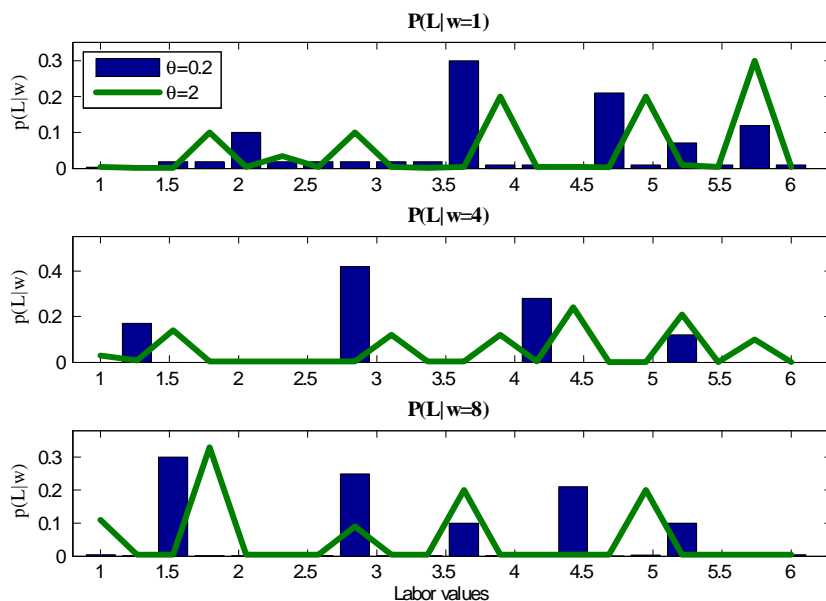
than the person with $\theta = 0.2$ does.[17]



**Figure 2b:** Optimal conditional distribution of labor for $\theta$=0.2 (bar) and $\theta$=2 (line)

## Model predictions

I define a model as a set $M = \{\theta, \gamma, \eta\}$ and I study the behavior of consumption, labor and wealth for $\gamma = 1$, $\eta = 1$ and $\theta \in (\{0.02\}, \{0.2\}, \{2\})$. I choose three values of $\theta$ as a proxy for three types of individuals that face 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 utility, a value of $\theta$ 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

---

[17]Additional results are in Appendix B in the journal's archive. In particular, tables $10a - 10b$ in Appendix 7 show that the expected values of consumption and labor for a person with $\theta = 0.2$ is $E(c) = 3.95$ with standard deviation $\sigma(c) = 1.10$ for consumption and $E(l) = 2.68$ and $\sigma(L)\,1.06$ for labor. A person with $\theta = 2$ enjoys on average $E(c) = 3.55$ units of consumption with standard deviation $\sigma(c) = 1.79$ and works an average of $E(l) = 3.05$ hours with $\sigma(l) = 1.52$. The average information flow is 1.08 bits for an individual with $\theta = 0.2$ and 0.73 bits for a person with $\theta = 2$.

value in the interval for the second choice of $\theta$. Similar reasoning leads to $\theta = 0.02$ as lower value of $\theta$ when $\theta$ takes up values in $(0.01, 0.05)$. For values of $\theta$ above 5, households acquire very little information about wealth and set consumption and labor basically constant. Also, values of $\theta$ below 0.01 deliver a solution close to the full information case.

Tables $10a - 10c$ in Appendix B present the moments of the time series for consumption, labor and wealth for each $\theta$ in $\theta \in (\{0.02\}, \{0.2\}, \{2\})$ as average across Monte Carlo simulations and simplex points.[18] For each $\theta$, mean, standard deviation, auto- and cross-correlations of the simulated series are calculated after taking averages of $7,000$ Monte Carlo runs and simplex points. Figure 8-11 in Appendix B shows the time series of the model-generated consumption, labor and wealth. The figure confirms that the persistence of these series is higher the higher the information cost is. The results are summarized in the following findings:

*For the model $M(\theta, 1, 1)$, $\theta \in (\{0.02\}, \{0.2\}, \{2\})$ :*

**Finding 1.** *the standard deviation of consumption over labor is greater than 1. The volatility is higher the higher the shadow cost of information is. Moreover, the higher the shadow cost of information, the more: (1) the more sluggish consumption and labor are; (2) sizeable the readjustment of both behavior is once changes in wealth are acknowledged.*

**Finding 2.** *the autocorrelation of consumption and labor is above 0.8.*

**Finding 3.** *the contemporaneous cross correlation between consumption and labor is above 0.80. Moreover, consumption and labor are more correlated to lagged values of wealth than contemporaneous values of wealth.*

*Finding 1* says that the model generates excess volatility of consumption with respect to labor supply. This result depends on the bound of information-processing capacity and

---

[18]Following the same methodology, tables $10d - 10f$ in Appendix B produce the statistics for the cases $M = (\{0.2, 3, 1\}, \{0.2, 1, 4\}, \{0.2, 1, .25\})$.

its interplay with the utility function. To illustrate this point, consider a consumer with log utility in consumption and Frisch labor supply elasticity equal to 1 receiving a positive shock to wealth. 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 cannot track wealth exactly but through signals on wealth as precise as his channel allows. He holds off changing consumption and keeps the same labor supply while he processes signals about wealth. If he has accumulated too much savings – due to the low information content of the period-by-period signal–, he has enough to afford a permanently higher steady state value of consumption than before. On average, the variation of consumption is bigger than the one of labor because of the utility specification and also the information processing constraint.[19] The intuition for *Finding 2* lies on the mechanism through which the consumer updates his knowledge of wealth, in (7). Each period he chooses a signal on wealth, decides consumption and labor on the basis of that signal and forms his posterior on wealth given his decisions. The higher the processing cost, the less informative the signal is. This in turn means that most of the update derives from the observations of past values of consumption and labor. From the previous example, if a positive shock to wealth is acknowledged with delays, then the behavioral responses to this shock will be delayed resulting in a strong autocorrelation between current and past values of consumption and labor. *Finding 3* says that the model predicts a strong comovement of labor and consumption. It also shows a strong correlation between contemporaneous consumption and wealth when information flow is high. The rationale for this finding is that these variables are related in two ways. The first is through the budget constraint (6) which is used to update the prior on wealth. The second is through the optimal policy of the consumer which is chosen to link tightly behavior (consumption and labor) and wealth.

---

[19]These effects occur because (1) utility specification commands a relatively stronger desire to smooth consumption than leisure; (2) the higher the shadow cost of using the channel is, the less informative is the signal that the agent processes, the longer he waits to change behaviors.

## Robustness

Tables $10d - 10f$ in Appendix B illustrate the relationships among risk aversion, $\gamma$, Frisch elasticity of substitution, $1/\eta$, and shadow cost of information, $\theta$. Table $10d$ shows that, for a given $\theta$ and elasticity of substitution, the higher the coefficient of risk aversion is, the higher the mean and the lower the variance of consumption are. An intuitive argument uses the example above: a risk averse person hit by a positive shock to wealth would not react to the shock immediately by changing consumption or hours worked. Once he processed information about wealth, he had accumulated enough savings to increase both consumption and leisure. The more risk averse the person is and the higher the shadow cost of using the channel is, the more he would want to process information before changing his behavior. Thus, shadow costs of processing information enhance precautionary savings. For a given $\theta$ and $\gamma$, a lower the Frisch elasticity of substitution (from $1/\eta = 4$ to $1/\eta = 0.25$ 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 Frisch elasticity increases the income effect over the substitution effect. However, the presence of information processing constraint induces a substitution effect that lessens 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 $1/\eta = 0.25$, since they are not certain that their wealth is actually increasing. When the collected information signals high wealth, labor supply suddenly decreases. The opposite occurs when consumers process 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 goes to infinity, consider $M(\theta) \equiv \{\theta; (\gamma \to 1, \eta = 0)\}$ with $\theta \in (\{0.02\}, \{0.2\}, \{2\})$. The first observation is that as the information cost increases, average consumption, labor and information flow decrease, while the standard deviations of these series increase. This is also true for wealth. An intuitive explanation of this result comes from the comparison with the rational expectation model. Under full information, log-utility $(\gamma \to 1)$ people prefer to smooth consumption. Moreover, with infinite Frisch

elasticity, ( i.e., $\eta = 0$) labor supply adjusts according to wealth fluctuations to accommodate consumption smoothing. Under finite information-processing capacity, rational households choose signals about wealth with the same purpose. If the shadow cost is low, $\theta = 0.02$, consumers choose a signal about wealth informative enough to allow them to use labor supply to smooth fluctuations in wealth and, in turn, consumption. By contrast, if the shadow cost is high, $\theta = 2$, 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 time path for wealth inherits the higher variance of consumption and labor and, on average, has higher mean than in the $\theta = 0.02$-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 3.

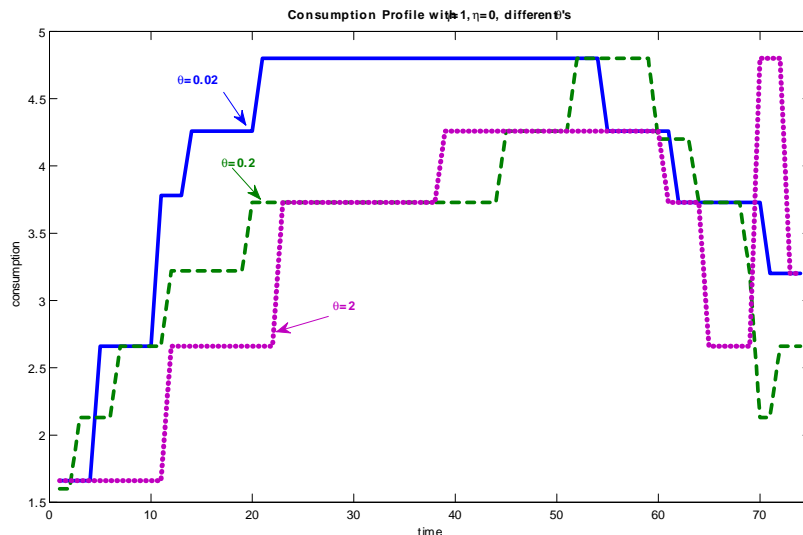Comparing consumption for $u(c,l)=\log(c)-\alpha l$



**Figure 3**: Sample path of consumption for different $\theta$'s.

The second observation is that consumption is smoother the lower information-processing shadow costs are. Figure 4 corroborates the argument. Consumers with $\theta = 0.02$ save at the beginning of the simulated period to enjoy high level of consumption later on.[20] By

---

[20]Recall that the time series of consumption, labor and wealth are derived from the model starting from a low value of wealth. Thus, at the beginning of the simulated periods people have very little savings.

contrast, consumers with $\theta = 2$ track wealth with noise 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. From Figure 4, note how individuals with less processing capacity $(\theta = 2)$ postpone 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 delay 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 preference specification, 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.
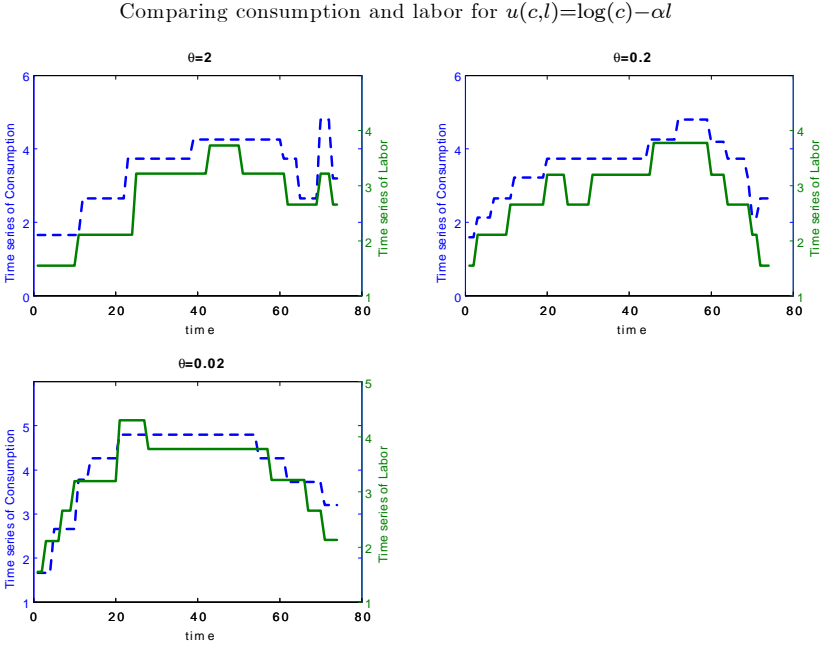


**Figure 4:** Time series of consumption and labor, various $\theta$

The third observation is that correlation between consumption and labor is higher the higher the information costs are. The intuition for this result is that the behavioral reactions to accumulation of wealth are delayed by individual's capacity of processing information. As the person has better knowledge of his wealth, he reviews both plans. People's actions are

mirrored in wealth accumulation. Individuals with $\theta = 0.02$ build up savings at the beginning of the period to spend gradually later on. This way of smoothing consumption in time is akin to that under full information.
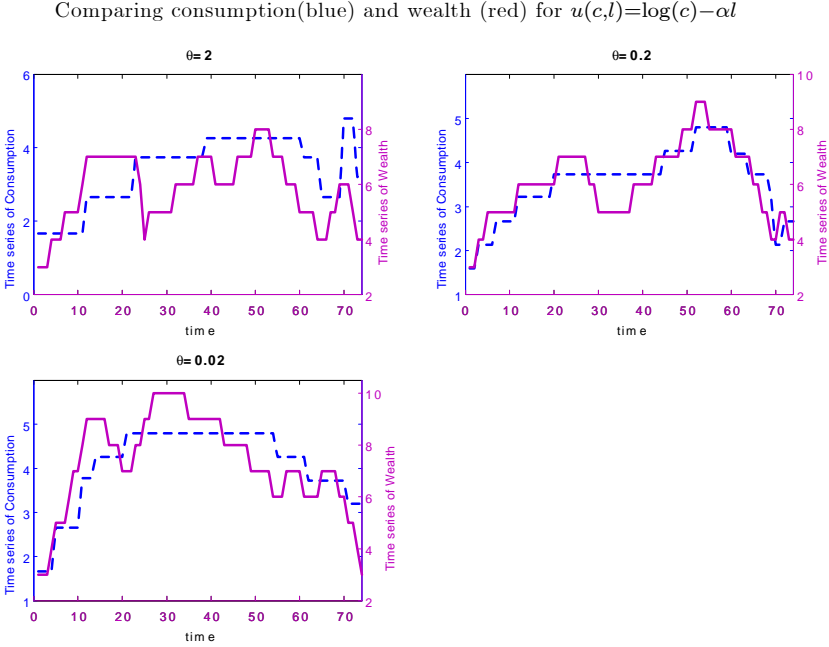
Comparing consumption(blue) and wealth (red) for $u(c,l)=\log(c)-\alpha l$



**Figure 5:** Time series of wealth and consumption, log-lin utility, various $\theta$

People with less processing capacity 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 revisions 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 0.47. A similar result holds for labor and lagged values of wealth. This result depends on the fact that a consumer with low capacity relies mostly on past value of consumption and labor to update his knowledge of wealth. While waiting, wealth accumulates and so does information until the consumer changes behavior. This mechanism implies lagged behavioral responses to wealth's fluctuations.

# Macroeconomic wedges and elasticities

This section presents the wedges (17)-(19) as a function of the shadow cost of information processing constraint, $\theta$. Throughout the section, I fix the specification of the utility to $\gamma \to 1$ and $\eta = 1$, that is $u(c, l) = \log(c) - (0.01) * l^2$.

I evaluate the intratemporal and intertemporal wedges by using data simulated from the solution of the problem with rational inattention and rational expectations. In particular, I solve the model for each $\theta$ where $\theta$ takes up values on a grid of 150 equi-spaced points in $[0, 3]$. Each $\theta > 0$ corresponds to a different model with rational inattention whereas $\theta = 0$ delivers the solution of the model under rational expectations.[21] Each of these models produces an optimal distribution from which I draw $7,000$ simulated time series of consumption, labor and wealth for $T = 100$ periods. The models of rational inattention and the model of rational expectations share the same discretization and calibration in Table 1 as well as the same paths for wages. The simulated time series from the rational expectation model and the rational inattention models are then fed to equations (17)-(19). I compute sample statistics averaging across the Monte Carlo runs, periods and, for the rational inattention models, simplex points.

I evaluate short-run and long-run elasticities by computing the behavioral responses of rational inattention agents to changes in wealth. Specifically, I shock the rationally inattentive economy with a temporary 10% increase in wage at the beginning of the simulation. The corresponding impulse-response function of consumptions and labor are then fed to the totally differentiated expressions (EC)-(EL) to evaluate short-run elasticities. Similarly, I assume a permanent 10% increase in wage at the beginning of the simulation and compute long-run elasticities by evaluating the totally differentiated expressions (S) and (6) through the impulse-response functions of consumption. Finally, I investigate what the estimates of long-run and short-run elasticities would be if an econometrician was to evaluate the totally

---

[21]More explicitly, I divide the interval $[0, 3]$ in 150 equally spaced points. Thus, I solve the problem (2) - (14) with 150 values of $\theta$ each of which corresponding to a different specification of the model $M(\theta) = \{\theta, \gamma \to 1, \eta = 1\}$ for $\theta \in [0, 3]$.

differentiated (S)-(EL) using data coming from a rationally inattentive economy.

## Intratemporal and intertemporal wedges

For $i = S, EC, EL$, figures 5 shows the mean, $E(\Lambda^i)$ – first column –, the standard deviation, $\sigma(\Lambda^i)$ – second column – and the signal to noise ratio, $E(\Lambda^i)/\sigma(\Lambda^i)$ – third column – of the expressions (17)-(19) as a function of $\theta$.

Consider the mean of the three wedges (first column). When $\theta = 0$ –i.e., when the capacity of the channel approaches infinity – the solution of the problem (2) - (14) mimics the solution of the rational expectation problem (11)-(5) resulting in no wedge, $E(\Lambda^i) = 0$. So long as $\theta$ increases, so does the mean of the wedges. For $\theta \in (0.2, 2)$, the static wedge increases at a lower rate than it does for $\theta \in (0, 0, 2)$. This change in slope reflects the change in the composition of $\Lambda^S$: Consumers with $\theta \in (0.6, 1.2)$ enjoy higher consumption and lower labor supply than consumers with $\theta \in (1.2, 2)$. Similarly, the slope of $\Lambda^{EC}$ and $\Lambda^{EL}$ flattens for $\theta \in (0.2, 2)$, reflecting slight revisions of consumption and labor plans as soon as consumers process more information about wealth. The same intuition explains the increase in wedges for $\theta \in (2, 2.5)$. For $\theta \geq 2.5$ consumers have little capacity to track wealth and to connect it to their consumption and labor. Thus, they set a constant level of consumption and work consistent with the static version of (6) even in the event of low income. These kinds of consumption and work behaviors allow consumers to process minimal information about wealth and still satisfying the budget constraint.

A comparison across means of the wedges reveals that the static wedge is bigger than the intertemporal wedges, reflecting the fact that finite processing capacity may generate sizeable long-run distortions. Moreover, the mean of $\Lambda^{EL}$ is higher than that of $\Lambda^{EC}$ for all $\theta \in (0, 0.25)$. The reason for this result lies in the specification of consumer's utility: with preference for smooth consumption relatively stronger than disutility of labor the agent offsets excessive fluctuations of wealth by changing his labor's schedule more than his consumption. With lower and lower capacities the revisions to both plans become infrequent

25

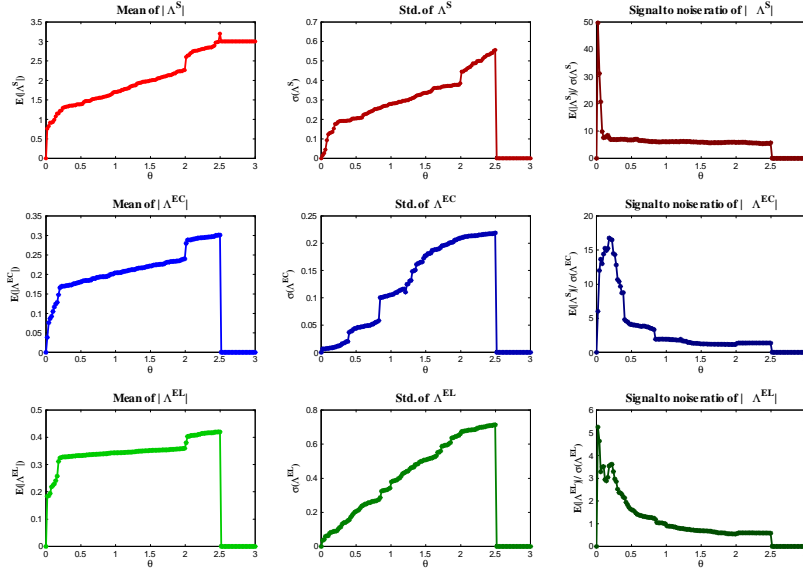and, when they occur, they represent a sharp deviation from previous plans.



**Figure 6:** Mean -first column-, standard deviation -second column- and SNR of $\Lambda^S$ (red), $\Lambda^{EC}$ (blue) and $\Lambda^{EL}$ (green).

Consider the standard deviation of the static wedge (2nd column of figure 6). For $\theta \in (0, 0.02)$, $\sigma\left(\Lambda^S\right)$ is very small and relatively flat so that the time series of consumption, labor and wealth mimic those of the solution under rational expectations. As $\theta$ increases, so does the standard deviation of $\Lambda^S$. In particular, the slope of $\sigma\left(\Lambda^S\right)$ gets steeper for $\theta \in (0.2, 2)$, and even more steeper in $\theta \in (2, 2.5)$. For $\theta \geq 2.5$, the consumer finds it too demanding to track wealth with high precision and resorts to a constant consumption and labor. Hence, in this case $\sigma\left(\Lambda^S\right) = 0$. The rationale for this behavior is as follows. For low values of $\theta$, there exists a high correlation of the time series of consumption, labor and wealth $(C, L, W(s))$ generated by fitting to (S) the simulated series from the rational inattention model. Such a strong correlation occurs because people with higher capacity (lower $\theta$) have better knowledge of the evolution of their wealth and its connection to consumption and labor supply than people with lower capacity have. Moreover, with better information

about wealth, the consumer has less volatile consumption and labor behavior. [22] Low variance of $C$ and $L$ and strong correlation of $(C, L, W(s))$ both reduce the variance of $\Lambda^S$. The consumer tracks wealth with low precision when $\theta$ takes up values in $(0.2, 2.5)$. As a result, his decisions $(C, L)$ are be more volatile and the correlation among $(C, L, W(s))$ increases the volatility of $\Lambda^S$. As noticed above, for $\theta \geq 2.5$, consumption and labor plans are essentially constant, resulting in a negligible $\sigma(\Lambda^S)$. A similar reasoning applies to the overall behavior of $\sigma(\Lambda^{EC})$ and $\sigma(\Lambda^{EL})$. By comparing the standard deviation of the three wedges it appears that for all $\theta$ except$\{\theta \neq 0, \theta \notin (2.5, 3)\}$, $\sigma(\Lambda^{EL})$ and $\sigma(\Lambda^S)$ dominate $\sigma(\Lambda^{EC})$. Higher short-run estimates of volatility of the *level* of labor with respect to the estimates of consumption come from the agent's preference for more stable consumption growth than labor growth.

The standard deviation of $\Lambda^{EL}$ is higher than that of $\Lambda^S$ for $\theta \in (0.5, 2.5)$. This result occurs because, for low values of $\theta$, the correlation among labor, consumption and wealth$(C, L, W(s))$ is higher than the correlation between wage growth and labor growth in $(\gamma^L(a, w(s)))$. As $\theta$ increases, the correlation between the wage growth rate and labor growth rate becomes weaker and the volatility of the growth of labor is higher than that of $(C, L, W(s))$ because of the delayed revisions in labor. I interpret this result as suggesting that, in this model, low information-processing capacity amplifies more the distortions to the intertemporal labor wedge than it does to the intratemporal labor wedge.[23]

Finally, consider the signal to noise ratio (SNR), $E(\Lambda^i)/\sigma(\Lambda^i)$, in the third column of figure 6, defined for $\theta \neq 0 \curlywedge \theta \notin [2.5, 3]$. The SNR provides a visual summary of the discussion of the mean and the variance of the wedges. The signal does not have a monotonic behavior since for $\theta \in (0.02, 0.25)$ the mean of the wedge increases more -on average- than the standard deviation of $\Lambda^i$, $i = S, EC, EL$. For $\theta \geq 25$, the noisiness of the wedges overpowers the strength of the signal up to the point $(\theta \to 2.5)$ in which the consumer is better off paying

---

[22]See Tables 10a-10d in Appendix B.

[23]This result seems at odds with Chari, Kehoe and McGrattan (2007). However, as noted by, e.g., Primiceri, Schaumburg and Tabalotti (2006), in a set up with two Euler equations, the data show that shocks to the intertemporal first order condition dominate those to the intra-temporal condition.

very little or no attention to wealth by planning constant consumption and leisure.

## Measures of Elasticities

Another assessment of the intertemporal and intratemporal conditions (S)-(EL) can be done by studying the responses of consumption and labor to changes in wages. In 4.2.1, I derive short-run elasticities by computing the reactions of consumption and labor to a temporary shock to wage. Similarly, I derive long-run elasticities through the average response of consumption and labor to a permanent shock to wage.

### Predictions from Impulse-Responses to shocks to wages

As stated in Mankiw, et. al. (1985), an intuitive way to think about elasticities is via impulse response functions. To gather a sense of short-run elasticities, one can look at the response of consumption and labor to a anticipated 10% increase in the mean of wage in $t = 0$ that lasts till $t = 1$. Likewise, the average impulse-response of consumption and labor to an anticipated and permanent 10% increase in the mean of wage might be a good proxy of the long-run elasticities.

Figures 7a-7b plot the response of labor and consumption to changes 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 7a displays the responses of consumption and labor to a permanent change in wage, while figure 7b shows the responses of consumption and labor to a temporary change in wage of the same amount. The top panels of figure 7a present the impulse response for aggregate consumption[24] (top right panel) and labor (top left panel). The bottom panels of figure 7a show responses of consumption (bottom right panel) and labor (bottom left panel) for types of consumers with three shadow costs of information-processing constraint, i.e., $\theta = 2$ (green

---

[24] *Aggregate* consumption, refers to the weighted average of consumption of types $\theta = 0.02$, $\theta = 0.2$ and $\theta = 2$ where each type gets the same weight.

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 .
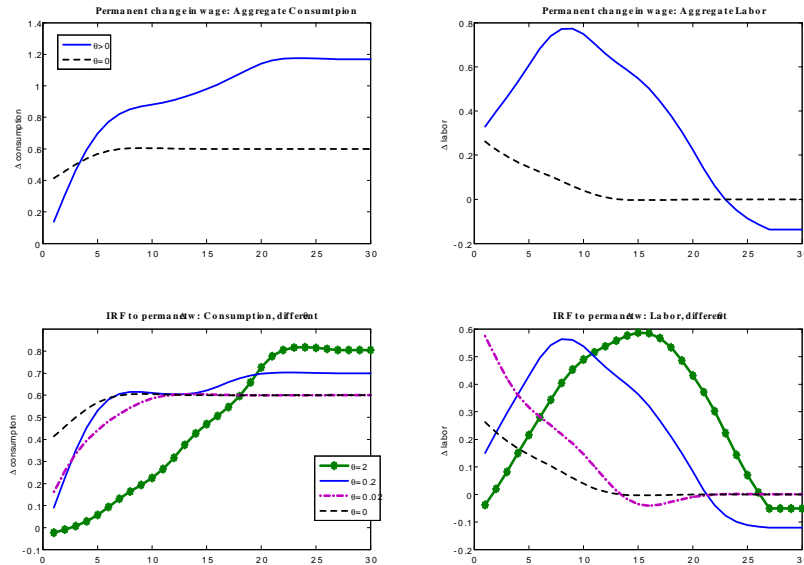


**Figure 7a.** Responses to a permanent 10%−change in wage: consumption (first column) and labor (second column)

Consider the impulse response of consumption and labor for different values of $\theta's$ (bottom panel). Under infinite information capacity, 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 steady state, the substitution effect commands a decrease in hours worked up until the household goes back to the work effort he exercised before the shocks. With shadow cost of processing information $\theta = 0.2$, people acknowledge the change in wealth due to the change in wage 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 is higher than before get strong by the increase in savings and in labor. Hence, $\theta = 0.2$-type reacts to those signals by moving consumption permanently up and slowly decreasing work effort. Note that the initial increase in labor 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 wage. 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 salary. 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), the model obtains that, when wages change permanently, long-run consumption increases 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 7b. It shows a temporary 10% change in wage. The first column of Figure 7b 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 7b plots the corresponding impulse-response for labor. In all the four pictures the black dashed line indicates impulse-responses under the infinite information solution.
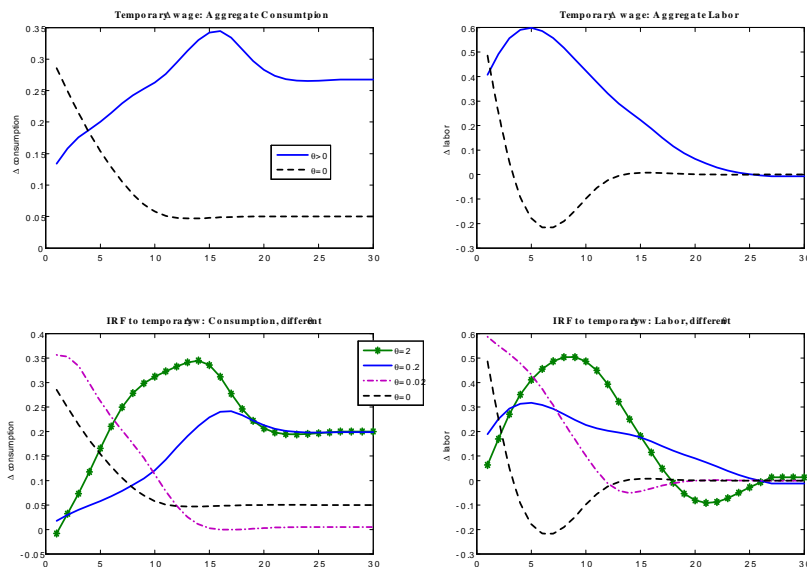


**Figure 7b.** Responses to a temporary 10%−change in wage: consumption (first column) and labor (second column)

The impulse-responses 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 a higher steady state value than it was before the shock. While consumption stabilizes to its new level, substitution effect prevails and type $\theta = 0.2$ trades off work for leisure up to the point in which labor reaches the same level as before the change in wage. People with $\theta = 0.02$ follows a similar pattern to people with infinite processing constraint. However, the effects of the temporary shock are more persistent than in the full information case. People with $\theta = 2$ adjust their consumption and labor decisions very slowly in response to the temporary shock. The logic of this result is akin to the one used for the permanent shock: income effect kicks in with delays 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. At this point, the substitution effect prevails and people start enjoying more leisure. The aggregate impulse responses (top panels) confirm these patterns. Finding 4 below summarizes the results of the impulse-response exercise:

**Finding 4.** *For the model* $M(\theta, 1, 1)$, $\theta \in (\{0.02\}, \{0.2\}, \{2\})$ :

    a. *The long-run effects of a permanent increase in wages are: (1) a significant in-crease in consumption, (2) a small or negative effect on hours worked. The effects on consumption are more pronounced the lower the information flow is.*

    b. *The short-run effects of a temporary increase in wages are: (1) a sluggish and positive response of labor (2) a response of consumption that is positive and more sluggish than labor. The effects on labor are more pronounced the lower the infor-mation flow is.*

The impulse-response functions suggest that a model of consumption- labor choices with

31

rational inattention is in principle consistent with the empirical estimation of long-run and short-run elasticities in U.S. data. Next section provides further support to this claim.

**Long-run and Short-run Elasticities**

Consider an econometrician who has data on consumption, labor and wealth generated by a rational inattention economy. Suppose he wants to measure elasticities of labor supply using a rational expectation model. Following Mankiw, et. al. (1985), to measure the short-run elasticities the econometrician evaluates the totally differentiated (EL),

$$\frac{dL_t}{ds_t} = \left( \frac{(\partial U \backslash \partial C_{t+1})}{s_{t+1}} \cdot \frac{dL_t \backslash L_t}{(\partial^2 U \backslash \partial L_t^2)} \right), \tag{20}$$

with data on $(C, L, W, s)$ drawn from the solution of (2) - (14). Similarly, he measures long-run elasticity of labor supply by totally differentiating (6) and (S),

$$\frac{dL_t}{ds_t} = \frac{s_t L_t \left( \partial^2 U \backslash \partial C_t^2 \right) + (\partial U / \partial C_t)}{- \left[ (s_t)^2 \left( \partial^2 U \backslash \partial C_t^2 \right) - \left( \partial^2 U \backslash \partial L_t^2 \right) \right]} \tag{21}$$

and by feeding (21) with the same simulated data.

Table 1 displays the sample statistics -mean and standard deviation- of short-run and long-run elasticities using (20)-(21) with averages of $7,000$ Monte Carlo runs of time series of consumption, labor and wealth drawn from the rational inattention solution $p_\theta^* (W; C, L)$, for $\theta \in (\{0\}, \{0.02\}, \{0.2\}, \{2\})$ and exogenously given wages.

Table 1 shows a disappointing performance of the rational expectations model (1)-(5), akin to the one reported in Mankiw, et. al. (1985): The elasticities generated by the optimality conditions (S)-(EL) evaluated with the artificial data from $p_\theta^* (W; C, L)$ imply an inelastic and backward-bending long-run supply curve whereas the short-run elasticities indicate strong substitution effect. From the intuition developed in section 4.2.1, rational inattention offers a way to reconcile the results in Table 1. In particular, when the person has very limited capacity $(\theta = 2)$, a temporary change in wage is not registered immediately

and as a result it is not reflected in behavioral changes. When the person realizes that his wealth has changed, he may want to adjust his work effort so that he can maintain a smooth consumption profile. Under rational expectations, the increase in labor following a temporary and, say, positive change in wage translates into a very elastic labor supply as shown in Table 2 (forth column, third row). Similarly, an individual with $\theta = 0.2$ who realizes with some delays the positive shift in wealth, may want to maintain his current labor effort and accumulate savings to be able to afford permanently higher levels of consumption and leisure than his pre-shock level. An econometrician that studies the data through the lens of rational expectation theory and produces Table 1 would reach similar conclusion as those of Mankiw, et. al. (1985) regarding the validity of stochastic dynamic optimization as a tool to interpret U.S. labor data.

## Conclusions

This paper compared the intertemporal and intratemporal conditions from the solution of a rational expectation model to the solution of a model with an information-processing constraint à la Sims. By evaluating agent's optimality conditions through consumption and labor behaviors arising from a rational inattention model and a rational expectation model, the paper showed that the model with finite processing capacity can simultaneously account for intertemporal and intratemporal wedges similar to those in the theoretical literature. Moreover, by evaluating the impulse responses of consumption and labor to shock to wages, the model with finite processing capacity can reconcile the estimates of long-run and short-run elasticities of substitution found in the empirical literature.

The two main implications of these findings are: (1) rational inattention theory provides a way to understand and reconcile puzzling results in the theoretical and empirical consumption-labor choices literature; (2) making a leap to a fully dynamic rational inattention consumption leisure problem produces results interesting enough to be worth of the

33

computational effort.

# Appendix A: Bellman Recursion and its properties[25]

## 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:

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

with

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

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 $HU$

$$HV\left(g\right) = H^{p_1}V\left(g\right),$$

$$HU\left(g\right) = H^{p_2}U\left(g\right).$$

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

---

[25] This appendix follows closely the work of Tutino (2009) to which I refer for details.

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

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

Hence,

$$||HV - HU|| =$$

$$||HV(g) - HU(g)|| =$$

$$H^{p_1}V(g) - H^{p_2}U(g) \le$$

$$H^{p_2}V(g) - H^{p_2}U(g) =$$

$$\beta \sum_{w \in W} \sum_{a \in A} \left[ (V^{p_2}(g'_a(\cdot))) - (U^{p_2}(g'_a(\cdot))) \right] p_2 g(w) \le$$

$$\beta \sum_{w \in W} \sum_{a \in A} (||V - U||) \, p_2 g(w) \le$$

$$\beta \, ||V - U||.$$

Recalling that $0 \le \beta < 1$ completes the proof. ∎

## 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 $V \le U \implies HV \le HU$, 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,

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

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

$\forall (g(w), c)$,

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

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

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

$$\le \sum_{w \in W} \left(\sum_{a \in A} u(c,l) p_1\right) \implies$$

$$H^{p_1} V(g) \le H^{p_1} U(g) \implies$$

$$H^{p_1} V(g) \le H^{p_2} U(g) \implies$$

$$HV(g) \le HU(g) \implies HV \le HU.$$

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

## 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 (12)-(15), 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 $\Gamma$ be the set that contains constraints (12)-(15) .For planning horizon $n = 0$, I have only to take into account the immediate expected rewards and thus I have that:

$$V_0\left(g\right) = \max_{p \in \Gamma} \left[ \sum_{w \in W} \left( \sum_{a \in A} u\left(c, l\right) p \right) g\left(w\right) \right] \tag{22}$$

and therefore if I define the vectors

$$\left\{ \alpha_0^i\left(w\right) \right\}_i \equiv \left( \sum_{a \in A} u\left(c, l\right) p \right)_{p \in \Gamma} \tag{23}$$

I have the desired

$$V_0\left(g\right) = \max_{\left\{\alpha_0^i(w)\right\}_i} \left\langle \alpha_0^i, g \right\rangle \tag{24}$$

where $\langle ., . \rangle$ denotes the inner product $\left\langle \alpha_0^i, g \right\rangle \equiv \sum_{w \in W} \alpha_0^i\left(w\right), g\left(w\right)$. For the general case, using equation (11):

$$V_n\left(g\right) = \max_{p \in \Gamma} \left[ \begin{array}{c} \sum_{w \in W} \left( \sum_{a \in A} u\left(c, l\right) p\left(c, l | w\right) \right) g\left(w\right) + \\ + \beta \sum_{w \in W} \sum_{a \in A} \left(V_{n-1}\left(g_a'\left(\cdot\right)_a\right)\right) p\left(c, l | w\right) g\left(w\right) \end{array} \right] \tag{25}$$

by the induction hypothesis

$$V_{n-1}\left(g\left(\cdot\right)|_a\right) = \max_{\left\{\alpha_{n-1}^i\right\}_i} \left\langle \alpha_{n-1}^i, g_a'\left(\cdot\right) \right\rangle \tag{26}$$

Plugging into the above equation (7) and by definition of $\langle ., . \rangle$ ,

$$V_{n-1}\left(g_a'\left(\cdot\right)\right) = \max_{\left\{\alpha_{n-1}^i\right\}_i} \sum_{w' \in W} \alpha_{n-1}^i\left(w'\right) \left( \sum_{w \in W} \sum_{a \in A} T\left(\cdot; w, c, l\right) \frac{\Pr\left(w, c, l\right)}{\Pr\left(c, l\right)} \right) \tag{27}$$

With the above:

$$
\begin{aligned}
V_n(g) &= \max_{p \in \Gamma} \left[ \begin{array}{l} \sum_{w \in W} \left( \sum_{a \in A} u(c,l) p \right) g(w) + \\ + \beta \max_{\{\alpha^i_{n-1}\}_i} \sum_{w' \in W} \alpha^i_{n-1}(w') \left( \sum_{w \in W} \left( \sum_{a \in A} \frac{T(\cdot; w, c, l)}{\Pr(c,l)} \cdot p \right) g(w) \right) \end{array} \right] \\
&= \max_{p \in \Gamma} \left[ \langle u(c,l) \cdot p, \ g(w) \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c,l)} \max_{\{\alpha^i_{n-1}\}_i} \left\langle \sum_{w' \in W} \alpha^i_{n-1}(w') T(\cdot; w, c, l) \cdot p, \ g \right\rangle \right]
\end{aligned}
\tag{28}
$$

At this point, it is possible to define

$$
\alpha^j_{p,a}(w) = \sum_{w' \in W} \alpha^i_{n-1}(w') T(\cdot : w, c, l) \cdot p.
\tag{29}
$$

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

$$
V_n(g) = \max_{p \in \Gamma} \left[ \langle u(c,l) \cdot p, \ g \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c,l)} \max_{\{\alpha^j_{p,a}\}_j} \left\langle \alpha^j_{p,a}, g \right\rangle \right],
\tag{30}
$$

and define:

$$
\alpha_{p,a,g} = \arg \max_{\{\alpha^j_{p,a}\}_j} \left\langle \alpha^j_{p,a}, g \right\rangle.
\tag{31}
$$

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

$$
\begin{aligned}
V_n(g) &= \max_{p \in \Gamma} \left[ \langle u(c,l) \cdot p, \ g \rangle + \beta \sum_{a \in A} \frac{1}{\Pr(c,l)} \langle \alpha_{p,a,g}, g \rangle \right] \\
&= \max_{p \in \Gamma} \left\langle u(c,l) \cdot + \beta \sum_{a \in A} \frac{1}{\Pr(c,l)} \alpha_{p,a,g}, \ g \right\rangle.
\end{aligned}
\tag{32}
$$

Now

$$
\{\alpha^i_n\}_i = \bigcup_{\forall g} \left\{ u(c,l) \cdot p + \beta \sum_{a \in A} \frac{1}{\Pr(c,l)} \alpha_{p,a,g} \right\}_{p \in \Gamma}
\tag{33}
$$

39

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

## .. 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 $\Gamma$, it is straightforward to see that through (23), 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 $M = \left| \{\alpha_{n-1}^j\} \right|$, 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. ∎

# References

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

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

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

[4] Blanchard, O. J., Galı', J., 2006. A New Keynesian Model with Unemployment. MIT Mimeo.

[5] Browning, M., Hansen L.P., Heckman, J. 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., Kim, S. 2005. On the aggregate labor supply, Federal Reserve Bank of Richmond Economic Quarterly, 91(1), 21-37.

[7] Chang, Y., Kim, S., 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., Kehoe, P. J., McGrattan, E. R. 2007. Business Cycle Accounting. Econometrica. 75 (3), 781–836.

[9] Chen, K., Imrohoroglu, A., Imrohoroglu, S., 2007. A Quantitative Assessment of the Decline in the U.S. Savings Rate. University of Southern California Mimeo.

[10] Davis, S. J., Faberman, R. J., Haltiwanger, J.C., 2006, The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links, Journal of Economic Perspectives 20, 3-26

[11] Gali', J., Gertler, M., Lopez-Salido, D.J., 2007. Markups, Gaps, and the Welfare Costs of Business Fluctuations. The Review of Economics and Statistics. 89 (1), 44–59.

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

[13] Gertler, M., Trigari, A., 2006. Unemployment Fluctuations With Staggered Nash Wage Bargaining. Journal of Political Economy, University of Chicago Press, 117(1), 38-86.

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

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

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

[17] Hansen, G. 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, S., Keane, M.P., 2004. Intertemporal Labor Supply and Human Capital Accumulation. International Economic Review. 45 (2), 601–641.

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

[21] Kimball, M., Shapiro, M., 2008, Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?. Mimeo, University of Michigan.

[22] Krusell, P., and Smith, A. Jr., 1998. Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy. 106 (5), 867-896.

[23] Lovejoy, W., 1989. Computationally feasible bound for partially observable Markov Decision problem, Operations Research, 39, 162-175,

[24] Lucas, R. E. Jr., Rapping, L. A., 1969. Real Wages, Employment, and Inflation. The Journal of Political Economy, 77 (5), 721–754.

[25] Mackowiak, B., Wiederholt, M., 2009. Optimal Sticky Prices under Rational Inattention, American Economic Review, 99 (3), 769-803.

[26] Mackowiak, B., Wiederholt, M., 2009. Business Cycle Dynamics under Rational Inattention, Mimeo.

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

[28] Mankiw, N. G., Rotemberg, J. J., Summers, L. H., 1985. Intertemporal substitution in Macroeconomics, The Quarterly Journal of Economics, 100 (1), 225-51.

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

[30] Mortensen, D. T., Pissarides, C.A., 1994. Job Creation and Job Destruction in the Theory of Unemployment. The Review of Economic Studies. 61 (3), 397–415.

[31] Mulligan, C. B., 2002. A Century of Labor-Leisure Distortions. NBER Working Paper 8774.

[32] Pistaferri, L., 2003. Anticipated and Unanticipated Wage Changes, Wage Risk and Intertemporal Labor Supply. Journal of Labor Economics 21, 729-754.

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

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

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

[36] Primiceri, G., Schaumburg, E., Tambalotti, A., 2006. Intertemporal Disturbances. NBER w.p. N0. 12243.

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

[38] Rogerson, R., 1988. Indivisible Labor, Lotteries and Eequilibrium. Journal of Monetary Economics. 21 (1), 3–16.

[39] Rogerson, R., Wallenius, J., 2007. Micro and Macro Elasticities in a Life Cycle Model with Taxes. Journal of Economic Theory, Elsevier, 144 (6), 2277-2292.

[40] Rotemberg, J. J., Woodford, M., 1991. Markups and the Business Cycle. NBER Macroeconomics Annual, 63–129.

[41] Shannon, C., 1948. A Mathematical Theory of Communication. Bell System Technical Journal, 27, 379–423 and 623–656.

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

[43] Sims, C. A., 1998. Stickiness. Carnegie-Rochester Conference Series On Public Policy, 49(1), 317–356.

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

[45] Sims, C. A., 2006 Rational Inattention: A research Agenda. American Economic Review Paper and Proceedings.

[46] Tutino, A., 2009. The Rigidity of Choices, Lifetime savings under information processing constraints. Princeton Universtiy, Mimeo.

[47] Uhlig, H., 2003. How Well do we Understand Business Cycles and Growth? Examining the Data with a Real Business Cycle Model. inWolfgang Franz, Hans Jürgen. Ramser, and Manfred Stadler (ed.), Empirische Wirtschaftsforschung: Methoden und Anwendungen, 32, 295–319.

[48] Verdú, S., 1996. The Exponential Distribution in Information Theory. Problems of Information Transmission. 32 (1), 86-95.

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

# Figure Legends

**Figure 1.** A typical day of a rationally inattentive person.

**Figure 2a.** Optimal conditional distribution of consumption for $\theta = 0.2$ (bar) and $\theta = 2$ (line).

**Figure 2b.** Optimal conditional distribution of labor for $\theta = 0.2$ (bar) and $\theta = 2$ (line).

**Figure 3.** Sample path of consumption for different $\theta$'s.

**Figure 4.** Time series of consumption and labor, various $\theta$.

**Figure 5.** Time series of wealth and consumption, log-lin utility, various $\theta$.

**Figure 6.** Mean -first column-, standard deviation -second column- and SNR of $\Lambda^S$ (red), $\Lambda^{EC}$ (blue) and $\Lambda^{EL}$ (green).

**Figure 7a.** Responses to a permanent $10\%$−change in wage: consumption (first column) and labor (second column).

**Figure 7b.** Responses to a permanent $10\%$−change in wage: consumption (first column) and labor (second column).

# Table

|  | $\theta = 0$ | $\theta = 0.02$ | $\theta = 0.2$ | $\theta = 2$ |
|---|---|---|---|---|
| *Short − run elasticities* |  |  |  |  |
| $\widehat{\eta}^{SR}$ | 0.91 | 0.43 | 0.14 | 0.09 |
| $\left(\hat{\sigma}_{\eta}^{SR}\right)$ | *(0.21)* | *(0.49)* | *(0.83)* | *(1.15)* |
| Implied $\hat{\varepsilon}^{SR}$ | 1.09 | 2.3 | 7.5 | 10.6 |
| *Long − run elasticities* |  |  |  |  |
| $\widehat{\eta}^{LR}$ | 1.12 | 2.8 | −12.1 | −6.1 |
| $\left(\hat{\sigma}_{\eta}^{LR}\right)$ | *(0.39)* | *(0.97)* | *(2.02)* | *(2.62)* |
| Implied $\hat{\varepsilon}^{LR}$ | 0.89 | 0.36 | −0.08 | −0.17 |

**Table 1:** Estimates of long-run and short-run elasticities from the RE model under RI data samples. Stdv. in parenthesis