Rational Inattention and Changes in Macroeconomic Volatility*

Jeta Menkulasi†
University of Maryland

November 9, 2009

Abstract

This paper offers a micro-founded theory of time variation in the volatility of aggregate economic activity based on rational inattention. I consider a dynamic general equilibrium model in which firms are limited in their ability to process information and allocate their limited attention across aggregate and idiosyncratic states. According to the model, a decrease in the volatility of aggregate shocks causes the firms optimally to allocate less attention to the aggregate environment. As a result, the firms' responses, and therefore the aggregate response, becomes less sensitive to aggregate shocks, amplifying the effect of the initial change in aggregate shock volatility. As an application, I use the model to explain the Great Moderation, the well-documented significant decline in aggregate volatility in the U.S. between 1984 and 2006. The exercise is disciplined by measurements of the changes in aggregate and idiosyncratic volatilities. The model can account for 90% of the observed decline in aggregate output volatility. 67% of the decline is due to the direct effect of the drop in the volatility of aggregate technology shocks and the other 23% captures the volatility amplification effect due to the optimal attention reallocation from aggregate to idiosyncratic shocks. A version of the model without rational inattention can capture the former effect but not the latter.

Keywords: rational inattention, business cycle, great moderation, aggregate TFP volatility, idiosyncratic TFP volatility

JEL Classification Numbers: E3, D8

*I am very grateful to Boragan Aruoba, John Haltiwanger and John Shea for their continued support and guidance. I also want to thank Antonella Tutino for her very helpful comments, conference and seminar participants at the Federal Reserve Board and University of Maryland.

†Jeta Menkulasi, University of Maryland at College Park, MD 20740. Phone: (240) 565-9187. E-mail: menkulasi@econ.umd.edu

1 Introduction

There was a well-documented decline in U.S. macroeconomic volatility lasting from the mid-1980s until 2006, followed by a renewed high macroeconomic volatility since 2007. This paper aims to explain the Great Moderation and to help understand the return to increased macroeconomic volatility.

During the Great Moderation, the volatility of aggregate output in the U.S. declined by 50%. The leading explanations of the Great Moderation include better monetary policy, structural changes such as better inventory management, and lower volatility of shocks hitting the economy. The first two explanations have proven to account only for part of the decline in macroeconomic volatility.[1]

As for the 'good luck' hypothesis, one can explain a 50% decline in output volatility in a standard RBC model only to the extent that the volatility of aggregate technology shocks declines by the same amount.[2] This opens the question of whether aggregate TFP volatility has in fact experienced such a decline. TFP series compiled by Basu, Fernald, and Kimball (2006) at an annual frequency covering the period 1949 - 1996 show only a 15% decline in the volatility of TFP innovations during the Great Moderation. Quarterly series by Fernald (2009) covering a longer time period 1949 - 2006 and using a different methodology exhibit a 34% decline.[3]

This clearly poses a problem for the 'good luck' hypothesis using a standard RBC model. If pure technology shocks have experienced at most a 34% decline in volatility, a RBC model can explain only a 34% decline in output volatility. This paper offers a mechanism that breaks this linear relationship between aggregate TFP shock volatility and output volatility. I propose an imperfect information setting in the form of rational inattention, in which changes in the volatility of aggregate shocks are amplified. Benchmark calibration of the model shows that a 34% decline in aggregate TFP shock volatility can generate a 46% decline in output volatility.

Rational inattention captures the idea that agents in the economy base their decisions not on the true state of the economy but on the perceived state, which is conditioned on their information set (Sims, 2003). Limited in their ability to process information, agents choose the optimal nature and precision of signals to reduce their uncertainty regarding the true state of the economy. One can think about the problem as a signal extraction problem, where the signal's noise properties are endogenously determined. In other words, the precision of the signals received as well as their

[1] Ahmed, Levin, and Wilson (2004), Arias, Hansen, and Ohanian (2006) and Stock and Watson (2003) compare hypotheses and conclude that in recent years the U.S. economy has to a large extent simply been hit by smaller shocks.

[2] See Arias, Hansen and Ohanian (2006) for a discussion of aggregate TFP volatility changes and the Great Moderation. Standard RBC models are characterized by an almost linear relationship between the volatility of aggregate technology shock and the volatility of aggregate output. This relationship is exactly linear up to a first order approximation and very close to linear for higher order appoximations.

[3] Basu, Fernald, and Kimball (2006) correct for aggregation issues, variable capacity utilization, deviations from constant returns to scale and imperfect competition. Fernald (2009) builds a quarterly series of total factor productivity that corrects only for variable capacity utilization.

statistical properties are choice variables. The restriction on the ability to process information limits how precise the signals can be. In the case where there is more than one state that agents in the economy are interested in tracking, the information processing problem becomes one of attention allocation: how to allocate information (attention) across multiple states, or in signal extraction terminology, how to allocate precision across multiple signals. This allocation will depend on the relevance of each state in the objective function as well as the properties of their stochastic processes, such as their relative persistence and volatility. More information will be allocated to variables with a higher variance or lower persistence for a given variance.[4]

This paper applies this 'attention allocation' problem to an otherwise standard RBC model with heterogenous firms and explores the transmission mechanism of shocks in the economy. The focus of the paper is the relationship between the volatility of aggregate technology and the volatility of aggregate outcomes such as output, labor, investment and consumption. Firms' profits depend on both aggregate and idiosyncratic state variables. Bounded in their ability to process information, they have to decide how to allocate the information flow across states. Given a higher relative volatility of the idiosyncratic state, firms will allocate more attention to the idiosyncratic environment and hence be more responsive to idiosyncratic shocks and less responsive to aggregate shocks. This leads to a dampening and delay in the response of endogenous variables to an innovation in the aggregate shock.

As the relative volatility of idiosyncratic versus aggregate states changes, so does the optimal allocation of attention. In the face of a decline in aggregate TFP shock volatility ('good luck', in the terminology of the Great Moderation literature), firms will reallocate their attention away from the aggregate environment since the relative volatility of the idiosyncratic environment has increased. This leads to an additional moderating effect. Hence, the decline in the volatility of aggregate outcomes is bigger than the decline in the volatility of the aggregate shock. This is in stark contrast with the full information version of the model, which is the standard rational expectations RBC model.

Evidence on plant-level data compiled by Davis, Haltiwanger, Jarmin and Miranda (2006) show that firm-level employment growth rate volatility has declined during the Great Moderation period by 9%, as compared to the 40-50% decline in its aggregate counterpart (Figure 1).[5]. Using indirect inference, I estimate a similar (9%) decline in the volatility of idiosyncratic TFP, which combined with the 34% decline in aggregate TFP volatility, this implies an increase in the idiosyncratic-to-aggregate volatility ratio.[6]

In the benchmark calibration this model can account for 90% of the decline in aggregate output volatility experienced by the U.S. in the past 30 years. 67% of the decline is due to direct effect of the drop in the volatility of aggregate technology shocks and the other 23% captures the volatility amplification effect due to the optimal attention reallocation from aggregate to idiosyncratic shocks.

[4] See Maćkowiak and Wiederholt (2009a)

[5] Figure 1 reports the 10-year window rolling standard deviations for firm-level and aggregate employment growth rates. The rolling standard deviations are normalized to 1 for the baseline year 1980.

[6] See Section 5.1 for details on the indirect inference exercise.

This paper presents the idea that the reduction in macroeconomic volatility in the mid-1980s has not been solely due to smaller aggregate shocks, but also to an increase in the relative volatility of idiosyncratic shocks as compared to aggregate shocks, which via an attention re-allocation has altered equilibrium behavior.

While I focus on the Great Moderation as the most obvious case study in the time variation of aggregate volatility, it is important to note that this mechanism is more general than the application in this paper. By allowing the idiosyncractic environment to play a role for aggregate dynamics, rational inattention in this model offers a new relationship between microeconomic and macroeconomic volatility. Because the idiosyncratic environment serves as a diversion of attention, changes in idiosyncratic volatility can affect aggregate dynamics without any change in the aggregate technology shock process. In order to expose the role of idiosyncratic shocks for aggregate dynamics more directly, I ask whether changes in the idiosyncratic state volatility *alone* can produce changes in aggregate volatility. My calibrated model shows that a hypothetical 25% increase in the volatility of the idiosyncratic state alone can produce an 11% decline in the volatility of aggregate output.

Starting with the financial crisis of 2007, there has been a renewed high degree of macroeconomic volatility. To the extent that there has been an increase in the volatility of the underlying aggregate shocks in the economy, this model predicts a reallocation of attention towards the aggregate environment by agents in the economy. This will in turn amplify initial changes in the volatility of aggregate shocks. Hence, the current increase in macroeconomic volatility might be partially due to more volatile aggregate shocks and partially due to more attention being reallocated towards the macroeconomic environment.

There have been several applications of rational inattention in the literature. Maćkowiak and Wiederholt (2009a) study the response of prices to aggregate nominal shocks versus idiosyncratic shocks in a partial equilibrium framework. They show how the attention allocation mechanism of firms under rational inattention leads to prices being more responsive to idiosyncratic shocks and less responsive to aggregate nominal shocks. This paper differs from Maćkowiak and Wiederholt (2009a) in two dimensions. First, I apply this mechanism in a general equilibrium real business cycle framework to study how rational inattention affects the transmission mechanism of aggregate technology shocks. Second, this paper discovers a new outcome of rational inattention, which is a volatility amplification effect. One main contribution of this paper is that I conduct a disciplined quantitative exercise of whether the Maćkowiak and Wiederholt (2009a) mechanism can explain the Great Moderation.

Applications of rational inattention in a dynamic general equilibrium setting include Paciello (2008), Luo and Young (2009) and Maćkowiak and Wiederholt (2009b). Paciello (2008) and Maćkowiak and Wiederholt (2009b) explore the differential response of prices to various aggregate and idiosyncratic shocks.[7] Rational inattention is shown to account for the sluggish response

[7] The main difference between Mackowiak and Widerholt (2009b) and Paciello (2008) and is that the latter considers only two aggregate shocks, whereas the former includes idiosyncratic shocks as well.

of prices to monetary shocks on one hand and their quicker adjustment to neutral technology shocks on the other.

Luo and Young (2009) introduce rational inattention in a stochastic growth model with permanent technology shocks and explore the extent to which rational inattention can enrich the weak internal propagation mechanism of shocks in RBC theory. This paper overlaps with theirs in that we both study the propagation mechanism of technology shocks in an RBC framework. It differs on the question of interest as well as in the solution method employed. I explore the second moment effects of rational inattention in an RBC framework, with the Great Moderation being the main case study. I also solve for a competitive equilibrium, which allows for a solution of rational inattention models with multiple state variables and accounts for general equilibrium effects on the propagation of shocks.

The paper is organized as follows: section 2 introduces the tools from information theory that are applied in my rational inattention setting. Section 3 introduces the benchmark model. In section 4, I study a simple version of the model that has an analytical solution to illustrate the main mechanism in the paper. Section 5 presents the calibration procedure and the numerical results for the benchmark model. In section 6, I distinguish between the roles of rational inattention (decision making under information processing constraints and one state variable) and attention allocation (rational inattention with multiple state variables). I show that simply restricting the ability to process information without having the problem of allocating information does not lead to a volatility amplification effect. Section 7 examines whether changes in the volatility of the idiosyncratic environment alone can lead to changes in aggregate volatility. Section 8 concludes.

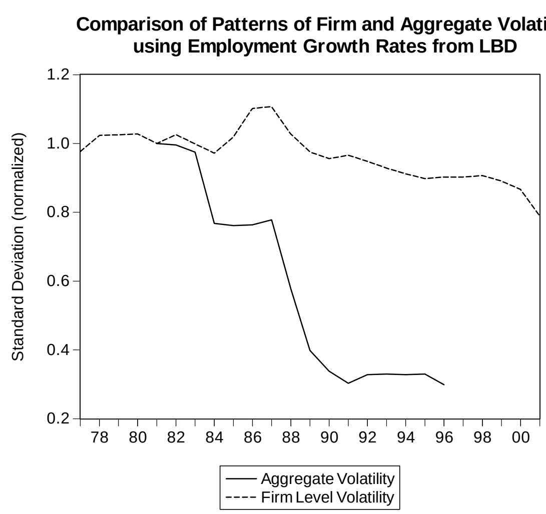

Figure 1. Source: Longitudinal Business Database (LBD), Davis, Haltiwanger, Jarmin and Miranda (2006).

2 Information Processing Constraints

In this section I introduce concepts from Information Theory that are used to quantify information flow and discuss how one can model a constraint in processing information. The rate of information flow is measured as the rate in uncertainty reduction, where the uncertainty regarding a random variable is measured by its *entropy*. Consider a random variable X, whose probability density function is $f(X)$. The entropy of X equals $-E\left[\log(f(X)\right]$. It's important to note that uncertainty about a random variable does not depend on its realizations but on the probability distribution of those realizations. Given the Gaussian setting of the model that will follow, I consider the entropy of a normally distributed variable. If X is normally distributed, then its entropy equals

$$H(X) = \frac{1}{2}\log_2(2\pi e Var(X))$$

Hence, the uncertainty regarding a normally distributed variable is summarized by its variance. *Conditional entropy* measures the conditional uncertainty of random variable X given another random variable Y. When X and Y follow a joint normal distribution, the conditional entropy becomes

$$H(X|Y) = \frac{1}{2}\log_2(2\pi e Var(X|Y))$$

Having quantified the uncertainty of a random variable, information flow is then defined as the rate at which this uncertainty is reduced. More specifically:

$$I(X;Y) = H(X) - H(X|Y)$$

That is, the rate of information flow between two random variables equals the difference between prior uncertainty and the posterior uncertainty. In the case that the two variables are independent from each other, the reduction in uncertainty will be zero, since knowing Y gives no information regarding X and hence the prior and posterior uncertainty will be the same. Constraints in the ability to process information are modelled as limits in the rate at which uncertainty about a random variable can be reduced. Formally, an information processing constraint is defined as:

$$I(X;Y) \leq \kappa$$

where κ is the capacity of the channel through which information is processed, which places an upper bound on the rate of uncertainty reduction through this channel. The channel is referred to as the device through which individuals process information (e.g. their brain) and the capacity refers to a technological constraint on the maximum amount of information that can be processed through this channel (Sims, 1988, 2003, 2006). As Sims (2006) notes, it's important to distinguish between various economic environments where such a description of uncertainty and limited information is logically consistent. Information processing constraints measured as limits to the capacity of a Shannon channel, as defined above, are consistent with an environment where information is publicly available and the only cost to making use of this information is the human information-processing capacity cost.

3 The model economy

In this section I develop a dynamic general equilibrium model representing an economy populated by households and firms. Given the availability of data on firm-level volatility, I will focus on the decision making process of firms facing a constraint in their information processing capabilities. There is a continuum of firms that produce a homogenous product using labor and capital and face a decreasing returns to scale production function as well as firm-specific technology shocks. Households are assumed to make their consumption, labor and investment decisions under perfect information. That is, they don't face constraints in their information processing capacity. This assumption is made for tractability purposes.

3.1 Firms

This part of the model is similar to Restuccia and Rogerson (2004) as well as Bartelsman (2006) with the main features of the model being diminishing returns to scale and heterogenous production units as in Hopenhayn (1992) and Hopenhayn and Rogerson (1993). The main difference between this model and the above papers is that I abstract from the entry and exit decision of firms.

The assumption of decreasing returns to scale allows me to pin down firm-level employment and capital, which will then form the basis of comparison with the firm-level dynamics we see in the data. There are two approaches to obtaining a non-degenerate distribution of firm size, the first being a single-good model where firms operate under decreasing returns to scale and perfect competition, and the second being a model with differentiated products and imperfect competition, which yields a non-degenerate distribution in size due to curvature in preferences. To avoid concerns about price setting and to keep the model as close as possible to the standard RBC model, I use decreasing returns to scale to get a non-degenerate distribution of firm size.[8]

The production technology each firm faces is

$$y_{it} = e^{a_t} e^{a_{it}} k_{it}^{\alpha} l_{it}^{\delta} \quad , \alpha + \delta < 1 \tag{1}$$

where a_t and a_{it} are the common and idiosyncratic components of firm-specific TFP respectively. In an environment of heterogeneous firms and decreasing returns to scale there may be a motive for entry and exit of firms. To avoid keeping track of this dimension I assume that in equilibrium there is no entry or exit. One can think of various institutional barriers that could make such movements very costly for a firm. In this model firms are not heterogenous in the products they produce but rather in the idiosyncratic TFP levels they face. They differ in their production levels as well as in

[8] A non-degenerate distribution of firm size is important in order to explore the role of the idiosyncratic environment.

the level of labor and capital they hire. Common and idiosyncratic components of firm-level TFP follow exogenous stochastic processes defined by

$$a_t = \rho_A a_{t-1} + \varepsilon_t \tag{2}$$

$$a_{it} = \rho_I a_{it-1} + u_t \tag{3}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $u_t \sim N(0, \sigma_u^2)$, and both variables are *iid* over time.

Profits in each period are

$$\pi(k_{it,} l_{it,} w_{t,} r_t) = e^{a_t} e^{a_{it}} k_{it}^{\alpha} l_{it}^{\delta} - w_t l_{it} - r_t k_{it} \tag{4}$$

where the wage and rental rate in the economy are taken as given by the firm.

The firm has to choose the level of capital and labor inputs that maximizes its profits subject to the informational constraints it faces. Formally firm i in period t chooses k_{it}^* and l_{it}^* to solve the following problem

$$\max_{\{k_{it}, l_{it}\}} \left[E \sum_{\tau=t}^{\infty} \tilde{\beta}_\tau \Pi(k_{i\tau,} l_{i\tau,} w_{\tau,} r_\tau, a_{\tau,} a_{i\tau}) | s_i^t \right]$$

where $s_i^t = \{s_{i,1,} s_{i,2,} ......s_{i,t}\}$ is the history of realizations of the signal process for firm i up until time t. The stochastic process of the signals that the firm chooses is an endogenous variable. Knowing how its signals affect its information set and hence its optimal input demand decisions, each firm chooses the precision of the signals it receives. The endogeneity of the signals' noise is the main difference between rational inattention in this model and signal extraction.[9] In order to ensure the stationarity of the attention allocation problem, I assume that the firm in at period 0 receives an infinite sequence of past signals $s_i^0 = \{s_{i,-\infty}, ...s_{i,-2,} s_{i,-1,} s_{i,0}\}$. Formally the problem of firm i in period 0 is

$$\max_{\{s_{it}\} \in S} E \left[\sum_{t=0}^{\infty} \tilde{\beta}_t \Pi(k_{it,}^* l_{it,}^* w_{t,} r_t, a_{t,} a_{it}) \right] \tag{5}$$

subject to

$$I(\{w_{t,} r_t, a_{t,} a_{it}\}; \{s_{it}\}) \leq \kappa \tag{6}$$

where $I(.)$ stands for the average flow of information between the states the firm is trying to track and the signals it chooses to receive regarding those states, and κ is the maximum amount of information the firms can process per period. Without any further constraints on the structure of signals, the problem that firms face in period 0 implies that firms choose the joint distribution of signals and state variables, which captures all the information signals contain about the state vector. This obviously makes the solution quite difficult due to the curse of dimensionality. To avoid such a problem I impose restrictions on the set of signals and take a quadratic approximation of the objective function to allow for a much easier solution to the firm's problem. I make the following assumptions on the set S. First, signals today do not contain any information about future shocks.

[9] See Sims (2003) for a discussion on signal extraction models and rational inattention.

Second, the vector of signals that a firm receives can be partitioned into a subset of signals regarding only the aggregate state (w_t, r_t, a_t) and another subset of signals regarding the idiosyncratic state a_{it}, so that $s_{it} = (s_{it}^A, s_{it}^I)'$, where $\{s_{it}^A, w_t, r_t, a_t\}, \{s_{it}^I, a_{it}\}$ are independent (this can be true only if $\{w_t, r_t, a_t\}, \{a_{it}\}$ are independent, which is assumed to be the case). The partition assumption implies that paying attention to the aggregate state and the idiosyncratic state are two separate activities. Third, $\{s_{it}^A, s_{it}^I, w_t, r_t, a_{t,}a_{it}\}$ follows a stationary Gaussian process. Gaussianity of the signals implies Gaussianity of the posterior distribution, which can be shown to be optimal when the optimization problem is quadratic (Sims, 2006). Given the tractability of a log-quadratic Gaussian (LQG) setting, I take a log-quadratic approximation of the objective function. The question of how good such an approximation is will be addressed in the calibration section of the paper. All the noise in the signals is assumed to be idiosyncratic, which is consistent with the idea that errors in tracking the state of the economy come from constraints in the ability to process information, not constraints in the availability of information (Sims 2003, 2006).[10]

The problem is set such that firms are assumed to choose the nature of their signals in period 0. This is not a restriction since it is optimal for the firm to choose its signal structure once and for all. Given the log-quadratic approximation of the profit function, the objective function of the firm will depend only on conditional variances. In addition, given the stationary Gaussian environment that the firms operate in, conditional variances are independent of realizations and constant over time. In period zero, the firm correctly anticipates future conditional variances and has no incentive to reallocate attention.[11]

3.1.1 Perfect Information

Before solving the imperfect information problem, I summarize the solution to the firm's problem under perfect information, which will be used in the attention allocation problem of each firm.

Proposition 1 *Under perfect information, that is, when firms perfectly observe* $\{a_t, a_{it}, w_t, r_t\}$ *every period, the log-linearized decision rules for the firm are*

$$\hat{l}_{it}^F = \frac{1}{1-\alpha-\delta}\left[a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t\right] \tag{7}$$

$$\hat{k}_{it}^F = \frac{1}{1-\alpha-\delta}\left[a_t + a_{it} - \delta\hat{w}_t - (1-\delta)\hat{r}_t\right] \tag{8}$$

and aggregate labor and capital follow

$$\hat{L}_t = \frac{1}{1-\alpha-\delta}\left[a_t - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t\right] \tag{9}$$

$$\hat{K}_t = \frac{1}{1-\alpha-\delta}\left[a_t - \delta\hat{w}_t - (1-\delta)\hat{r}_t\right] \tag{10}$$

Proof. See Appendix F. ■

[10] The above mentioned assumptions also appear in Maćkowiak and Wiederholt (2009a,b) and Paciello (2007).

[11] See Maćkowiak and Wiederholt (2009a)

It is important to emphasize that under perfect information, the aggregate economy looks exactly like the representative agent RBC model with DTRS technology on firms' side, where the aggregates depend only on aggregate technology shocks and idiosyncratic shocks disappear. Solving for the full-information equilibrium is important in drawing out the main differences rational inattention introduces to aggregate behavior, which are that idiosyncratic volatility matters for aggregate behavior and that aggregate volatility responds more than one-for-one to a change in the volatility of aggregate TFP.

3.1.2 Rational Inattention

I start by taking a log-quadratic approximation of the profit function expressed in terms of log deviations from steady state. Denoting $\hat{\pi}(a_t, a_{it}, \hat{k}_{it}, \hat{l}_{it}, \hat{w}_t, \hat{r}_t) = \pi(e^{a_t}, e^{a_{it}}, \bar{K}e^{\hat{k}_{it}}, \bar{L}e^{\hat{l}_{it}}, \bar{w}e^{\hat{w}_t}, \bar{r}e^{\hat{r}_t})$, where bars denote steady state values and carats denote percentage deviations from steady state, the second order Taylor approximation of $\hat{\pi}$ around (0,0,0,0,0,0) is given by

$$\tilde{\pi}(a_t, a_{it}, \hat{k}_{it}, \hat{l}_{it}, \hat{w}_t, \hat{r}_t) \simeq \hat{\pi}(0,0,0,0,0,0,0) + \hat{\pi}_1 a_t + \hat{\pi}_2 a_{it} + \hat{\pi}_3 \hat{k}_{it} + \hat{\pi}_4 \hat{l}_{it} + \hat{\pi}_5 \hat{w}_t + \hat{\pi}_6 \hat{r}_t$$

$$+ \frac{\hat{\pi}_{11}}{2} a_t^2 + \frac{\hat{\pi}_{22}}{2} a_{it}^2 + \frac{\hat{\pi}_{33}}{2} \hat{k}_{it}^2 + \frac{\hat{\pi}_{44}}{2} \hat{l}_{it}^2 + \frac{\hat{\pi}_{55}}{2} \hat{w}_t + \frac{\hat{\pi}_{66}}{2} \hat{r}_t$$

$$+ \hat{\pi}_{12} a_t a_{it} + \hat{\pi}_{13} a_t \hat{k}_{it} + \hat{\pi}_{14} a_t \hat{l}_{it} + \hat{\pi}_{15} a_t \hat{w}_t + \hat{\pi}_{16} a_t \hat{r}_t$$

$$+ \hat{\pi}_{23} a_{it} \hat{k}_{it} + \hat{\pi}_{24} a_{it} \hat{l}_{it} + \hat{\pi}_{25} a_{it} \hat{w}_t + \hat{\pi}_{26} a_{it} \hat{r}_t$$

$$+ \hat{\pi}_{34} \hat{k}_{it} \hat{l}_{it} + \hat{\pi}_{35} \hat{k}_{it} \hat{w}_t + \hat{\pi}_{36} \hat{k}_{it} \hat{r}_t + \hat{\pi}_{45} \hat{l}_{it} \hat{w}_t + \hat{\pi}_{46} \hat{l}_{it} \hat{r}_t + \hat{\pi}_{56} \hat{w}_t \hat{r}_t$$

Using the approximated profit function, the optimal capital and labor inputs that the firm chooses are

$$\hat{l}_{it}^* = \phi_a^L E[a_t|s_i^t] + \phi_I^L E[a_{it}|s_i^t] + \phi_w^L E[w_t|s_i^t] + \phi_r^L E[r_t|s_i^t] \tag{11}$$

$$\hat{k}_{it}^* = \phi_a^K E[a_t|s_i^t] + \phi_I^K E[a_{it}|s_i^t] + \phi_w^K E[w_t|s_i^t] + \phi_r^K E[r_t|s_i^t] \tag{12}$$

where $\{k_{it}^*, l_{it}^*\}$ stand for optimal capital and labor input under rational inattention.[12]

For comparison the solution of firm i in period t under full information is:

$$\hat{l}_{it}^F = \phi_a^L a_t + \phi_I^L a_{it} + \phi_w^L w_t + \phi_r^L r_t \tag{13}$$

$$\hat{k}_{it}^F = \phi_a^K a_t + \phi_I^K a_{it} + \phi_w^K w_t + \phi_r^K r_t \tag{14}$$

[12] Coefficients in the capital and labor input choices are as follows: $\phi_a^L = (\frac{\pi_{34}\pi_{13}}{\pi_{33}} - \pi_{14})$, $\phi_I^L = (\frac{\pi_{34}\pi_{23}}{\pi_{33}} - \pi_{24})$, $\phi_w^L = (\frac{\pi_{34}\pi_{35}}{\pi_{33}} - \pi_{45})$, $\phi_r^L = (\frac{\pi_{34}\pi_{36}}{\pi_{33}} - \pi_{46})$, $\phi_a^K = \frac{\pi_{34}}{\pi_{33}}(\frac{\pi_{34}\pi_{13}}{\pi_{33}} - \pi_{14}) - \frac{\pi_{13}}{\pi_{33}}$, $\phi_I^K = \frac{\pi_{34}}{\pi_{33}}(\frac{\pi_{34}\pi_{23}}{\pi_{33}} - \pi_{24}) - \frac{\pi_{23}}{\pi_{33}}$, $\phi_w^K = \frac{\pi_{34}}{\pi_{33}}(\frac{\pi_{34}\pi_{35}}{\pi_{33}} - \pi_{45}) - \frac{\pi_{35}}{\pi_{33}}$, and $\phi_r^K = \frac{\pi_{34}}{\pi_{33}}(\frac{\pi_{34}\pi_{36}}{\pi_{33}} - \pi_{46}) - \frac{\pi_{36}}{\pi_{33}}$. Equations (13) and (14) are identical to equations (7) and (8).

where $\{k_{it}^F, l_{it}^F\}$ stand for the optimal choices of labor and capital under full-information. As one can see from the equations above, $\hat{l}_{it}^* = E\left[\hat{l}_{it}^F|s_i^t\right]$ and $\hat{k}_{it}^* = E\left[\hat{k}_{it}^F|s_i^t\right]$. A firm operating under imperfect information chooses inputs on the basis of the perceived states $(E[a_t|s_i^t], E[a_{it}|s_i^t])$, whereas a firm operating under full information chooses inputs on the basis of the actual state (a_t, a_{it}). Anytime the input choices differ from those prevalent under full information, there is a loss in profits. This loss can be measured by subtracting from $\hat{\pi}(a_t, a_{it}, \hat{k}_{it}^*, \hat{l}_{it}^*, \hat{w}_t, \hat{r}_t)$ the equivalent expression under full information $\hat{\pi}(a_t, a_{it}, \hat{k}_{it}^F, \hat{l}_{it}^F, \hat{w}_t, \hat{r}_t)$, which simplifies the attention allocation problem without affecting the solution since the perfect information profits are independent of the signal choice.

The loss function is given by

$$L \equiv \hat{\pi}(a_t, a_{it}, \hat{k}_{it}^*, \hat{l}_{it}^*, \hat{w}_t, \hat{r}_t) - \hat{\pi}(a_t, a_{it}, \hat{k}_{it}^F, \hat{l}_{it}^F, \hat{w}_t, \hat{r}_t)$$

which can be simplified to

$$L = \frac{\hat{\pi}_{33}}{2}(\hat{k}_{it}^* - \hat{k}_{it}^F)^2 + \frac{\hat{\pi}_{44}}{2}(\hat{l}_{it}^* - \hat{l}_{it}^F)^2 + \hat{\pi}_{34}(\hat{k}_{it}^* - \hat{k}_{it}^F)(\hat{l}_{it}^* - \hat{l}_{it}^F)$$

using (13), (14) and the fact that $\hat{\pi}_3 = \hat{\pi}_4 = 0$. Here $\hat{\pi}_{44} = \bar{Y}\delta^2 - \bar{w}\bar{L}$, $\hat{\pi}_{33} = \bar{Y}\alpha^2 - \bar{r}\bar{K}$ and $\hat{\pi}_{34} = \bar{Y}\alpha\delta$. The first term of the loss function measures the loss in profits due to the suboptimal capital choice, whereas the second term measures the loss due to suboptimal labor decision. The last term in captures how the mistake in one variable affects the cost of a mistake in the other variable.

The attention allocation problem can now be stated as

$$\min_{\{s_{it}\}} E\left\{\sum_{t=0}^{\infty} \beta^t \left[\frac{\hat{\pi}_{33}}{2}(\hat{k}_{it}^* - \hat{k}_{it}^F)^2 + \frac{\hat{\pi}_{44}}{2}(\hat{l}_{it}^* - \hat{l}_{it}^F)^2 + \hat{\pi}_{34}(\hat{k}_{it}^* - \hat{k}_{it}^F)(\hat{l}_{it}^* - \hat{l}_{it}^F)\right]\right\} \tag{15}$$

subject to

$$\hat{l}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t) \tag{16}$$

$$\hat{k}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - \delta_t\hat{w}_t - (1-\delta)\hat{r}_t) \tag{17}$$

$$\hat{l}_{it}^* = E\left[\hat{l}_{it}^F|s_i^t\right] \tag{18}$$

$$\hat{k}_{it}^* = E\left[\hat{k}_{it}^F|s_i^t\right] \tag{19}$$

$$I(\{w_{t,}r_t, a_{t,}a_{it}\}; \{s_{it}\}) \leq \kappa \tag{20}$$

The result that the input choices under rational inattention are linear projections of the optimal choices under perfect information is due to the objective function being quadratic. Given the assumption that signals regarding idiosyncratic and aggregate states are orthogonal, the information flow can be expressed as the sum of information flow that aggregate signals reveal for aggregate states, and the information flow that idiosyncratic signals reveal for idiosyncratic states. Formally,

$$I(\{w_t, r_t, a_t, a_{it}\}; \{s_{it}\}) = I(\{w_t, r_t, a_t\}; \{s_{it}^A\}) + I(\{a_{it}\}; \{s_{it}^I\})$$

where s_{it}^A and s_{it}^I represent the set of signals regarding the aggregate and idiosyncratic states respectively. In this model there is only one idiosyncratic state whose true realization firms would like to track, namely the idiosyncratic component in firm-level TFP. On the other hand there are multiple aggregate states that firms are interested in tracking. In the multiple state case there is an additional constraint that needs to be satisfied

$$\Omega_A \succeq \Omega_{A|S^A}$$

where Ω_A is the prior variance-covariance matrix of the aggregate state vector and $\Omega_{A|S^A}$ is the posterior variance-covariance of the same aggregate vector conditional on the set of signals received. That is, the difference between the prior and posterior variance-covariance matrix must be positive semi-definite. This constraint is otherwise called the *non-subsidization* constraint, which places a restriction on the precision of signals. Without this constraint, the decision-maker can improve the precision of one signal by erasing information (forgetting) about another variable (which can be achieved without violating the constraint on information processing capacity, equation (20). One can think of this condition as a type of irreversibility constraint on the amount of information acquired about a particular state variable. Further details on how information flow is derived can be found in appendix D.

3.2 Households

The household sector is represented by a representative consumer which has access to perfect information and a complete set of Arrow Debreu contingent securities. By perfect information I mean that the household knows the whole history of the relevant states including period t realizations. Households maximize expected discounted utility given by

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \theta \frac{L_t^{1+\psi}}{1+\psi} \right]$$

where C_t is aggregate (average) consumption, L_t is the household's supply of labor, γ is the coefficient of relative risk aversion, ψ is the inverse labor supply elasticity and θ captures the level of disutility of labor. Households make their decisions subject to the following budget constraint

$$C_t + K_{t+1} = w_t L_t + (1 + r_t - d) K_t + \Pi_t \tag{21}$$

where w_t and r_t are the wage and rental rate respectively, d is the depreciation rate of capital and Π_t is the dividend yield from households' ownership of firms. Labor is assumed to be homogeneous.

The transversality condition is

$$\lim_{T\to\infty} E_0[\pi_{t=0}^{T}(1+r_{t+1})^{-1}]K_{T+1} = 0 \tag{22}$$

Knowing the history of $\{w_t, r_t\}$ including the period t realization, households choose period t's consumption, labor supply and next period's capital holdings $\{C_t, K_{t+1}, L_t\}$. First order conditions obtained from the household's problem are as follows

$$C_t^{-\gamma} w_t = \theta L_t^{\psi} \tag{23}$$

$$C_t^{-\gamma} = \beta E[C_{t+1}^{-\gamma}(1+r_{t+1}-d)] \tag{24}$$

3.3 Equilibrium

The set of conditions to be satisfied in equilibrium include first order conditions for the household problem

$$C_t^{-\gamma} w_t = \theta L_t^{\psi} \tag{25}$$

$$C_t^{-\gamma} = \beta E[C_{t+1}^{-\gamma}(1+r_{t+1}-d)] \tag{26}$$

the resource constraint:

$$C_t + K_{t+1} - (1-d)K_t = Y_t \tag{27}$$

labor and capital market equilibrium, where the prevalent wage and rental rate are determined

$$L^s(w_t, r_t, a_t) = \int_i L^d(s_i^t) \tag{28}$$

$$K^s(w_t, r_t, a_t) = \int_i K^d(s_i^t) \tag{29}$$

market clearing condition $Y_t = \int y_{it} di$; and processes for the aggregate and idiosyncratic components of firm-level TFP, which are assumed to follow an AR(1) with parameters to be calibrated using aggregate and firm specific data sets

$$a_{it} = \rho_I a_{it-1} + u_{it} \tag{30}$$

$$a_t = \rho_A a_t + \varepsilon_t \tag{31}$$

$$\int a_{it} di = 0$$

4 Special Case: No Capital and White Noise Disturbances

In order to illustrate the main mechanism in the model here I solve a labor-only version of the incomplete information model where disturbances follow a white noise process. The main differences from the benchmark case are that households cannot save, the production function is $y_{it} = e^{a_t} e^{a_{it}} l_{it}^{\delta}$, and the model is static. Such a setting allows for an analytic solution, which clarifies the main mechanism in the paper.

4.1 Full Information

The equilibrium amount of aggregate hours employed in production, the wage rate and the level of consumption in the economy under full information are

$$\hat{L}_t^F = \frac{1-\gamma}{1+\psi-\delta+\delta\gamma} a_t \tag{32}$$

$$\hat{w}_t^F = \frac{\psi+\gamma}{1+\psi-\delta+\delta\gamma} a_t \tag{33}$$

$$\hat{C}_t^F = \frac{1+\psi}{1+\psi-\delta+\delta\gamma} a_t \tag{34}$$

The solution under full information shows, once again, that the aggregate variables in the economy are determined only by the aggregate component of TFP and that no characteristic of the idiosyncratic environment matters for aggregate dynamics. In the following subsection it will be analytically shown how macroeconomic dynamics under rational inattention do depend on the idiosyncratic environment and how this leads to a volatility amplification effect.

4.2 Attention Allocation Problem

In this section I assume that the common and idiosyncratic components of firm-level TFP follow Gaussian white noise processes with respective variances σ_a and σ_{ai}.

Each firm's attention allocation problem becomes

$$\min_{\{s_{it}\}} E\left[\sum_{t=0}^{\infty} \beta^t \frac{\pi_{33}}{2} (\hat{l}_{it}^* - \hat{l}_{it}^F)^2\right] \tag{35}$$

subject to

$$\hat{l}_{it}^F = \frac{1}{1-\delta}\left(a_t + a_{it} - \hat{w}_t\right) \tag{36}$$

$$\hat{l}_{it}^* = \frac{1}{1-\delta}\left(E\left[a_t|s_i^t\right] + E\left[a_{it}|s_i^t\right] - E\left[\hat{w}_t|s_i^t\right]\right) \tag{37}$$

$$I\left(\{\hat{w}_t, a_{t,} a_{it}\}; \{s_{it}\}\right) \leq \kappa \tag{38}$$

There are three variables of interest to the firms, namely the aggregate and idiosyncratic component of TFP as well as the average wage in the economy. I start with the guess that in equilibrium the wage rate satisfies $w = \varphi a_t$ and solve the attention allocation problem as a function of such a guess. Instead of tracking three variables, the firms track only the aggregate and idiosyncratic component of TFP. Given the guess on the wage rate, which linearly depends on the realization of the aggregate TFP a_t, tracking $\{\hat{w}_t, a_{t,} a_{it}\}$ is the same problem as tracking $\{a_{t,} a_{it}\}$.

Given the quadratic nature of the objective function and the Gaussian white noise process assumed for the states, one can prove that the optimal signals that firms choose take the form of "the true state + white noise".[13] Hence, we have

$$\begin{aligned} s_{1it} &= a_t + u_{it} && (39) \\ s_{2it} &= a_{it} + \varepsilon_{it} && (40) \end{aligned}$$

where $u_{it} \sim N(0, \sigma_u^2)$ and $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$.

After receiving the signals regarding the two exogenous states, firms form their posteriors using Bayes' Rule

$$E(a_t | s_{1it}) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_u^2} s_{1it}$$

$$E(a_{it} | s_{2it}) = \frac{\sigma_{a_i}^2}{\sigma_{a_i}^2 + \sigma_\varepsilon^2} s_{2it}$$

These posteriors are substituted in the firm's objective function and the attention allocation problem becomes

$$\min_{\frac{\sigma_a^2}{\sigma_u^2}, \frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}} \left\{ \left(\frac{1-\varphi}{1-\delta} \right)^2 \sigma_a^2 \left[\left(\frac{1}{\frac{\sigma_a^2}{\sigma_u^2} + 1} \right)^2 + \frac{1}{\frac{\sigma_a^2}{\sigma_u^2}} \right] + \left(\frac{1}{1-\delta} \right)^2 \sigma_{a_i}^2 \left[\left(\frac{1}{\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2} + 1} \right)^2 + \frac{1}{\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}} \right] \right\} \tag{41}$$

subject to

$$\frac{1}{2} \log_2 \left(1 + \frac{\sigma_a^2}{\sigma_u^2} \right) + \frac{1}{2} \log_2 \left(1 + \frac{\sigma_{ai}^2}{\sigma_\varepsilon^2} \right) \leq \kappa \tag{42}$$

Each firm minimizes its losses due to imperfect information by choosing the signal-to-noise ratios $\{\frac{\sigma_a^2}{\sigma_u^2}, \frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}\}$.

[13] Given that a_t is assumed to follow a white noise process , w_t is also white noise with a variance of $\phi^2 \sigma_a^2$.

Optimal signal-to-noise ratios for each signal are

$$\frac{\sigma_a^2}{\sigma_u^2} = \begin{cases} 0 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \leq 2^{-2\kappa} \\ (1-\varphi)\frac{\sigma_a}{\sigma_{ai}} 2^{\kappa} - 1 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \in (2^{-2\kappa}, 2^{2\kappa}) \\ 2^{2\kappa} - 1 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \geq 2^{2\kappa} \end{cases}$$

$$\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2} = \begin{cases} 2^{2\kappa} - 1 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \leq 2^{-2\kappa} \\ \frac{2^{\kappa}}{(1-\varphi)\frac{\sigma_a}{\sigma_{ai}}} - 1 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \in (2^{-2\kappa}, 2^{2\kappa}) \\ 0 & \text{if } (1-\varphi)^2 \frac{\sigma_a^2}{\sigma_{ai}^2} \geq 2^{2\kappa} \end{cases}$$

For each signal there are two possible corner solutions: one in which the firm chooses to allocate no attention (information flow) at all and one where it chooses to allocate all of the attention at its disposal. Zero information flow allocated to a signal implies that the signal-to-noise ratio of that signal is zero. That is, the firm chooses to receive an infinitely noisy signal regarding that particular state. When a particular signal receives all of the information flow, its signal-to-noise ratio represents the maximum precision that the signal can have given the limits on the ability to process information.

The guess regarding the average wage rate in the economy implies a guess regarding the average equilibrium labor employed in the economy via the general equilibrium effects from the household equilibrium conditions. Hence we have

$$L_t = \frac{\varphi - \gamma}{\psi + \gamma\delta} a_t \tag{43}$$

as the implied guess for aggregate labor.

Using the results above, I solve for the fixed point, in which the aggregate response of labor to aggregate shocks equals the initial guess (43)

$$\varphi^* = \begin{cases} \gamma & \text{if } \sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}} > (1-\gamma)2^{\kappa} \\ \frac{\psi+\gamma}{1+\psi-\delta+\delta\gamma}\left(1 - \frac{\psi+\gamma\delta}{\psi+\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}} 2^{-\kappa}\right) & \text{if } \sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}} < \frac{(1-\delta-\gamma+\gamma\delta)2^{-\kappa}}{1-\delta+(\psi+\gamma\delta)(1-2^{-2\kappa})} \\ \frac{(\psi+\gamma\delta)(1-2^{-2\kappa})+\gamma(1-\delta)}{1-\delta+(\psi+\gamma\delta)(1-2^{-2\kappa})} & \text{if } \sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}} < \frac{(1-\delta-\gamma+\gamma\delta)2^{-\kappa}}{1-\delta+(\psi+\gamma\delta)(1-2^{-2\kappa})} \end{cases}$$

Using the assumptions on the signals and the derived information flow constraint, the *interior* solution to the attention allocation problem is as follows[14]

$$\hat{L}_t = \hat{L}_t^F \left(1 - \frac{1}{1-\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}} 2^{-\kappa}\right) \tag{44}$$

[14] See appendix for details on these derivations.

$$\hat{C}_t = \hat{C}_t^F \left(1 - \frac{\delta}{1+\psi}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}}2^{-\kappa}\right) \tag{45}$$

$$\hat{w}_t = \hat{w}_t^F \left(1 - \frac{\psi+\gamma\delta}{\psi+\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}}2^{-\kappa}\right) \tag{46}$$

where $\{L_t^F, C_t^F, w_t^F\}$ are the full information solutions for labor, consumption and wage rate respectively, as defined in equations (32), (33) and (34).

The solution under rational inattention differs from the full information solution in two important ways. First, rational inattention leads to dampened responses of all aggregate variables to a change in aggregate TFP. Second, the responses of all aggregate variables to an innovation in aggregate TFP are a function of aggregate and idiosyncratic TFP volatility. The latter is the key result of this model. Endogeneizing the information set in a rational inattention sense introduces a first-order effect of aggregate and idiosyncratic shock volatilities. The key parameter for this result is the relative volatility of idiosyncratic to aggregate shocks, σ_{ai}^2/σ_a^2. As this ratio increases, idiosyncratic TFP is relatively more volatile compared to aggregate TFP, which leads to a reallocation of attention (information flow) towards the idiosyncratic state at the cost of less attention being allocated to the aggregate state. The less information allocated to aggregate TFP, the stronger the dampening of the responses of macroeconomic aggregates to an aggregate TFP shock. It is important to note that even though the model is solved using log-linearization methods, endogeneizing the information set leads to a first-order effect of aggregate and idiosyncratic TFP volatilities on the impulse responses of endogenous variables. In this way I can isolate the second-moment effect on equilibrium outcomes originating only from the imperfect information part of the model. The result that the response of macroeconomic variables to aggregate TFP is a function of relative volatility leads to another result, which I will call the *volatility amplification* effect. A 1% change in aggregate TFP volatility leads to more than a 1% change in the volatility of macroeconomic aggregates. A standard RBC model solved using higher order approximations to account for potential second-moment effects has almost no volatility amplification, i.e. a 1% change in aggregate TFP volatility leads to an approximately 1% change in macroeconomic volatility. Hence, the two main results that imperfect information in the form of rational inattention delivers are a dampening in the response of all macroeconomic aggregates to an innovation in aggregate TFP, and an amplification in the response of macroeconomic volatility to a change in aggregate TFP volatility. The first result is the usual result of imperfect information settings. Inability to see the true state of the economy with no error leads to a smoother response and potentially a delay, as shown below in the numerical solution for more generalized stochastic processes. The amplification in volatility occurs because a decline in the volatility of the aggregate TFP shock has the direct effect of lowering the volatility of the aggregate outcome, as well as the indirect effect of inducing agents to pay less attention to aggregate shocks, leading to an additional moderating effect.

In order to see this amplification effect analytically, I compute the elasticity of each aggregate variable's volatility with respect to the volatility in aggregate TFP: $\epsilon_{\sigma_a^2}^{var(X)} = \left(\frac{\partial var(X)}{\partial \sigma_\varepsilon^2}\right)\left(\frac{\sigma_\varepsilon^2}{var(X)}\right)$. The volatility elasticities for each *aggregate variable* with respect to σ_a^2 are

$$\epsilon_{\sigma_a^2}^{var(Y)} = \frac{1}{1 - \frac{\delta}{1+\psi}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}2^{-\kappa}}} > 1 \tag{47}$$

$$\epsilon_{\sigma_a^2}^{var(L)} = \frac{1}{1 - \frac{1}{1-\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}2^{-\kappa}}} > 1 \tag{48}$$

These elasticities are the main concern of this paper. In the white noise case, this amplification effect is determined by the *relative volatility* of the idiosyncratic versus the common component of TFP, the information processing capacity, the risk aversion coefficient, the degree of decreasing returns to scale and the elasticity of labor supply. As the relative volatility increases, more attention is allocated to the idiosyncratic state, and firm-level actions respond less to aggregate states. This leads to a higher volatility amplification. As the capacity to process information increases, the more the economy moves towards full-information since more capacity is available to allocate to each state. Hence, the higher the information processing capacity, the lower the volatility amplification.

In order to explain the relationship between behavioral and technological parameters affecting volatility amplification, I run the following thought experiment: suppose the economy experiences a decrease in the volatility of the common component of TFP. On the labor demand side of the economy, that is firms, the fall in aggregate volatility will lead to a reallocation of attention away from the aggregate states and towards the idiosyncratic state. This in turn will lead firms to respond less to aggregate shocks. After aggregating all firms' responses, this leads to a lower volatility in aggregate labor demand. On the supply side of the labor market, that is households, a fall in the volatility of the common component of TFP will lead to a decline in labor supply volatility. Given that the labor market must be in equilibrium, the change in volatilities for labor demand and supply of labor must be the same. This implies that wage volatility must change in equilibrium. This change in wage volatility introduces general equilibrium effects in the attention allocation problem. One can show that for a risk aversion coefficient less than one ($\gamma < 1$), the higher the CRRA, the bigger the change in wage volatility required to restore labor market equilibrium for any given change in common TFP volatility. In this experiment, the bigger the fall in wage volatility, the bigger the fall in the volatility of the aggregate state that each firm wants to track. Hence, there is another round of attention reallocation in favor of idiosyncratic variables and the same process repeats itself. To see how the labor supply elasticity and returns to scale affect aggregate output volatility, one can use the following equations governing household labor supply and the resource constraint: $Y_t = a_t + \delta L_t = a_t + \frac{\delta}{\psi}(W_t - \gamma C_t)$, so that a given change in wage volatility will lead to higher changes in the volatility of output the closer production technology is to constant returns to scale (higher δ) and the higher the Frisch elasticity of labor supply (lower ψ). Thus, higher δ and lower ψ increase the volatility amplification effect.

At the unique interior solution the optimal amount of information allocated to the aggregate shock is

$$\kappa^A = \frac{1}{2}\log_2\left[\frac{1}{1-\delta+\psi+\gamma\delta}\left(\frac{\sigma_a}{\sigma_{a_i}}(1-\delta)(1-\gamma)+\ 2^{-\kappa}(\psi+\gamma\delta)\right)\right]+\frac{1}{2}\kappa \tag{49}$$

and the amount of information allocated to the idiosyncratic state is:

$$\kappa^I = \kappa - \kappa^A \tag{50}$$

Equation (49) shows that the amount of attention allocated to each variable depends on preference and technology parameters as well as the ratio of aggregate versus idiosyncratic volatility. Below I consider an experiment, designed to mimic the Great Moderation, in which preference and technology parameters do not change over time, while changes in the volatility of each shock affect the allocation of attention across states.

5 Numerical Solution of the Benchmark Model

This section provides the numerical solution to the benchmark model with serially correlated shocks presented in Section 3, which is a dynamic stochastic general equilibrium model similar to the standard RBC model with the exception of rational inattention on the part of firms. I explore how accounting for an endogenous information set affects the transmission mechanism of aggregate technology shocks to the economy.

5.1 Calibration

The period in the model is set to one quarter. Parameters that govern preferences and production technology are calibrated such that they match long-run values of postwar US aggregates. I follow standard calibration procedure as explained in Cooley and Prescott (1995) and Prescott (1986). Using steady state equations, β is chosen to match an annual real rate of return of 4%, which implies a value of 0.99 for β. The depreciation rate of 0.02 fixes the investment to capital ratio. Choosing a value of 1 for the coefficient of relative risk aversion reconciles the long-run observations for the US economy of constant per-capita leisure and steadily increasing real wages (Cooley, 1995).

There has been an extensive empirical literature trying to estimate the curvature of the profit function, which captures the decreasing returns to scale in the production function. Important papers include Thomas (2002), Thomas and Khan (2007), Cooper and Haltiwanger (2005), Fuentes, Gilchrist and Rysman (2006), and Hennessy and Whited (2005). The estimated curvature ranges from 0.5 to 0.9. In the benchmark model I follow Thomas and Khan (2007) and set the labor share to 0.64 and capital share to 0.245.

The parameter ψ determining the inverse of the Frisch elasticity of labor is set at 0.1 following Gali et al. (2005), who takes this value from micro estimates of the elasticity of labor supply with respect to the real wage. The parameter controlling the level of disutility of labor θ is then chosen such that households spend 1/3 of their time working.[15]. Parameters governing the persistence

[15] This number comes from microeconomic evidence on time allocation studies, such as Ghez and Becker (1975).

and standard deviation of the aggregate TFP shock are obtained using the quarterly series on TFP computed by Fernald (2007). I fit equation (31) to the detrended data for both the pre and post-1984 periods and obtain an autocorrelation coefficient of 0.98 for both periods and standard deviations of 0.0092 for 1960-1983 and 0.006 for 1984-2005 respectively. This implies a 34% decline in the volatility of the innovations in aggregate TFP and a 15% decline in the volatility of TFP itself.

5.1.1 Idiosyncratic TFP process

I use the evidence on plant-level data compiled by Cooper and Haltiwanger (2006) and Davis, Haltiwanger, Jarmin and Miranda (2006) to determine the parameters governing the process of firm-level productivity. There is only one moment in the model that can be exactly matched to the data and that is the standard deviation of firm-level employment growth rate. On the other hand, assuming an AR(1) process for the idiosyncratic TFP process, there are two parameters to be pinned down: the autocorrelation coefficient and the standard deviation. Given that both parameters cannot be pinned down, I fix the persistence parameter to different values and compute the implied standard deviation for the TFP process by matching the model's implications to the data.

There is little consensus on the persistence of idiosyncratic TFP shocks. Ideally this parameter should be estimated using firm-level panel data accounting for both common and idiosyncratic components to firm-level TFP. Unfortunately, no annual firm-level data set with the information needed to compute TFP is available for the US. The best available persistence estimates can be found in Cooper, Haltiwanger and Wallis (2007), who explicitly model a common and idiosyncratic component to firm-level shocks and estimate idiosyncratic level shock parameters indirectly by matching various moments in the data. Given different model specifications, the persistence parameters vary from 0.33 to 0.89.

I match the model's predictions for firm-level employment dynamics with moments from firm-level employment growth rate data provided by Cooper and Haltiwanger (2006) and Davis, Haltiwanger, Jarmin and Miranda (2006). The moments available from these studies are 10-year window rolling standard deviations of firm-level employment growth rates. The firm-level data in these studies is annual, whereas my model economy is quarterly. I aggregate the model to an annual frequency and obtain the firm-level growth rate in employment. Given the log-linearized version of the model and the additive form of the first order conditions, I can exactly pin down the volatility parameter of the idiosyncratic TFP process once I make an assumption on the persistence of the idiosyncratic TFP. The indirect inference exercise is done using the full-information version of the model. Inferring the parameters of the idiosyncratic process assuming perfect information has two advantages. First, it saves computational time and second, equilibrium firm-level responses to idiosyncratic shocks under rational inattention match almost perfectly the behavior under perfect information, since firms under my benchmark calibration optimally allocate close to 95% of their information flow to tracking the idiosyncratic state.

The first order condition with respect to labor for firm i in the full information model is as follows

$$L_{it} = \frac{1}{1-\alpha-\delta}\left[a_t - (1-\alpha)w_t - \alpha r_t + a_{it}\right]$$

where a_t is the aggregate TFP shock, whose parameters I take as given from Fernald (2007), and a_{it} is the idiosyncratic TFP.

Under full information, the equilibrium behavior of w_t and r_t is independent of the idiosyncratic TFP. Assuming aggregate and idiosyncratic TFP are AR(1) processes, their dynamics can be expressed as MA(∞): $a_t = \rho_A a_{t-1} + \varepsilon_t$ can be represented as $a_t = a^A(L)\varepsilon_t$ and $a_{it} = \rho_I a_{it-1} + u_{it}$ can be represented as $a_t = a^I(L)u_{it}$, where lag polynomials $a^I(L)$ and $a^A(L)$ are functions of their respective auto-correlation coefficients. As a result, the model's decision rules can also be expressed as MA processes, which yields the following representation of the first order condition above

$$L_{it} = \frac{1}{1-\alpha-\delta}\left[a^A(L)\varepsilon_t - (1-\alpha)W(L)\varepsilon_t - \alpha R(L)\varepsilon_t + a(L)u_{it}\right]$$

There are two unknown parameters in this decision rule, namely the persistence and standard deviation of the idiosyncratic TFP process. Given that the only firm-level moment available to me is the standard deviation of firm-level employment growth rate, I experiment with different persistence parameters suggested from the literature and then back out the implied standard deviation.

The firm-level data are in the form of 10-year window rolling standard deviations of firm-level employment growth rates

$$\sigma_{it} = \left(\frac{1}{10}\sum_{s=-4}^{5}\left(g_{it+s} - \bar{g}_i\right)^2\right)^{1/2}$$

where g_{it} is the firm-level growth rate in employment and $\bar{g}_i$ is its 10-year average. I compute the model-equivalent measure and calculate the implied idiosyncratic TFP volatility. For each sub-period (before and after 1984), I simulate the model 100 times with each simulation consisting of 300 periods. I then aggregate the model to an annual frequency and compute a time-series of the rolling standard deviation for the firm-level employment growth rate. I average the 10-year window rolling standard deviation for each sub-period and compute the implied idiosyncratic TFP. Table 1 reports the implied idiosyncratic standard deviation as well as the implied ratio of idiosyncratic-to-aggregate volatility for different assumed persistence parameters for the idiosyncratic shock.

Table 1. Implied standard deviation for the Idiosyncratic TFP shock

	pre 1984	post 1984	% change
Average standard deviation (firm-level employment growth rate data)	0.4996	0.4730	-9.46
Idiosyncratic TFP persistence $\rho_I = 0.95$			
Implied σ_u	0.1746	0.1653	-9.46
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	19.036	27.510	44.51
Idiosyncratic TFP persistence $\rho_I = 0.5$			
Implied σ_u	0.1537	0.1456	-9.47
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	16.763	24.226	44.52
Idiosyncratic TFP persistence $\rho_I = 0.3$			
Implied σ_u	.1435	.1359	-9.47
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	15.645	22.610	44.52

The results show that in order to match the annual data on firm-level volatility, the implied standard deviation for innovations of idiosyncratic TFP prior to 1984 ranges between 0.15 and 0.17, which is 15-19 times higher than the standard deviation for aggregate TFP for the pre-1984 period. The implied standard deviation for the post-1984 era ranges between 0.13 to 0.16, which is 22-25 times than that of aggregate TFP over this period. The ratio of idiosyncratic-to-aggregate TFP volatility has increased, despite a decline in both idiosyncratic and aggregate TFP volatility, because the decline in aggregate TFP volatility has been substantially higher than that of idiosyncratic TFP. This is the key stylized fact that will enable the calibrated model with rational inattention to generate a volatility amplification effect when applied to the Great Moderation episode. For the benchmark model below, I choose the persistence parameter for the idiosyncratic TFP process to be equal to that of the aggregate TFP process, $\rho_I = 0.95$. By setting the persistence parameter equal across the two processes I can focus on the relative volatility ratio as the main variable that determines the allocation of attention.

5.1.2 Calibrating the upper bound on information flow κ

The value of κ , the maximum information processing capacity, has implications for the per period loss of profits for each firm due to imperfect tracking of state variables as well as for the marginal value of information. As Sims (2003, 2006) shows, the Log-Quadratic-Gaussian setting is a good approximation when the marginal value of information flow is low and a bad approximation when the marginal value of information flow is high. Hence, κ is chosen in such a way as to imply a low marginal value of information. More specifically, as in Maćkowiak and Wiederholt (2009a,b), one can fix the marginal value of information and let κ be determined endogenously, or fix κ and let the marginal value of information be determined within the model. In both cases the marginal value of information must be a reasonably low number. I pick the latter strategy, because my goal is to evaluate the effect of changes in the stochastic processes of underlying shocks keeping fixed the information processing technology. In the benchmark calibration, $\kappa = 4.7$ bits, which implies a marginal value of information of 0.04% of a firm's steady state output and an expected per-period loss in profits of 0.07% of a firm's steady state output. I think these are reasonably low numbers.

Table 2 summarizes the benchmark calibration.

Table 2. Benchmark Parameters

Parameter	**Description**	**Values**
β	discount factor	0.99
γ	coefficient of relative risk aversion	1
ψ	the inverse of labor supply elasticity	0.1
d	depreciation rate	0.02
α	capital's share in output	0.256
δ	labor's share in output	0.64
θ	the level of disutility of labor	2.95
κ	upper bound on information flow (bits)	4.7
ρ_A	persistence parameter for aggregate TFP process	0.95
ρ_I	persistence parameter for idiosyncratic TFP process	0.95
σ_ε (pre-1984)	standard deviation of the innovation in aggregate TFP	0.0092
σ_ε (post-1984)	standard deviation of the innovation in aggregate TFP	0.006
σ_u (pre-1984)	standard deviation of the innovation in idiosyncratic TFP	0.1746
σ_u (post-1984)	standard deviation of the innovation in idiosyncartic TFP	0.1653

5.2 Results

Figure 2 displays impulse responses of aggregate variables to a one standard deviation positive shock to aggregate TFP under perfect information and rational inattention. All impulse responses presented in the paper represent percentage deviations from the nonstochastic steady state. For a given volatility of aggregate TFP, rational inattention leads to a dampening and delay in the responses of output, labor, consumption and investment to an innovation in aggregate TFP as compared to perfect information. This is due to a combination of reasons. First, agents in the economy are limited in their ability to process information, which implies imperfect tracking of the true state vector in the economy. The degree of this imperfection depends on how tight the information capacity constraint is. The tighter the constraint, the less precise the signals and the more dampening and delay will be observed. Existing studies on RBC models with rational inattention (e.g. Luo and Young 2009) have found significant departures from perfect information outcomes for a very low maximum bound on information flow (around .30 bits per time period, which is a quarter). In this model, a low information flow devoted to tracking the aggregate shock is an optimal outcome, which is the second explanation for the findings in Figure 1. Agents in this economy are endowed with 4.7 bits per period of information flow, but they optimally choose to allocate only 5% of this information flow to aggregate conditions. Hence, with most of the information flow allocated to the idiosyncratic environment, agents in the economy have a smooth and delayed response to an innovation in aggregate TFP.

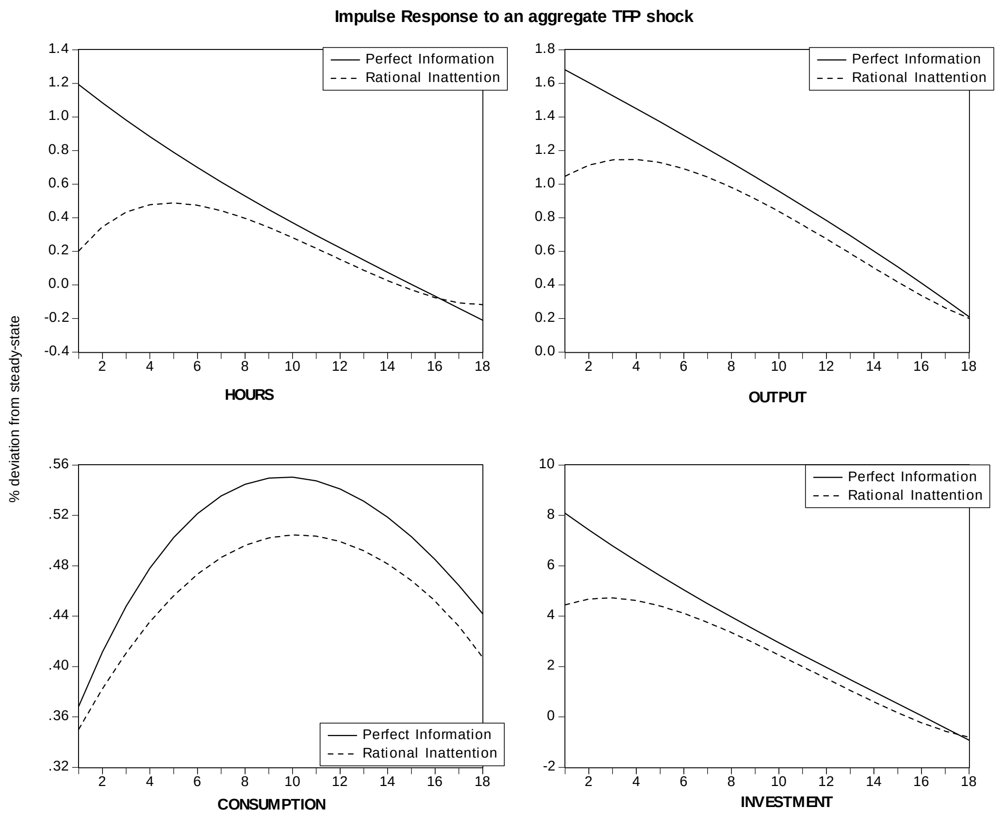

Figure 2. Impulse Response to an aggregate TFP shock.

Because firms optimally devote most of their attention to idiosyncratic outcomes, their response to idiosyncratic shocks under rational inattention is almost identical to that under perfect information, as shown in Figure 3. Labor and capital inputs are affected equally by the idiosyncratic shock. Hence, the impulse responses for both labor and capital to an innovation in the idiosyncratic TFP shock will be the same.

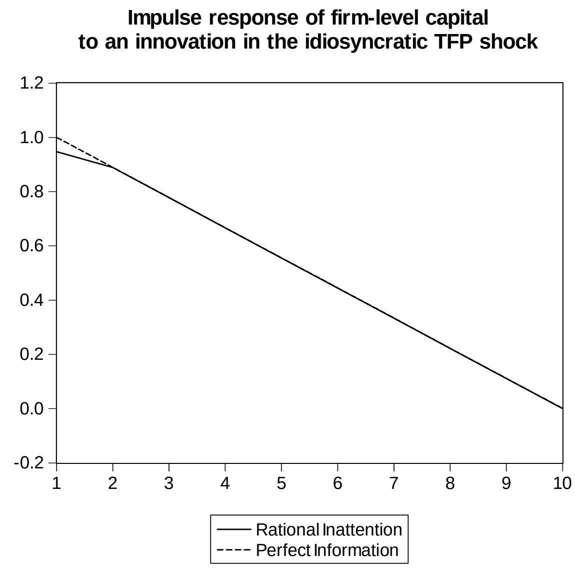

Figure 3. Impulse response of firm level input choices to an innovation in idiosyncratic TFP

Next I calculate the second moments implied by the benchmark model using the pre-1984 estimated aggregate and idiosyncratic TFP volatilities. I simulate the model 200 times, with each simulation consisting of 300 periods. I apply the HP filter to the simulated data and compute the moments presented in Table 3. Major differences between the perfect information and rational inattention models are observed in the volatility of aggregate variables. Note that given the simplifying assumption that the household sector in the economy has full information, there is little difference in the volatility of consumption. However, investment, hours and output are markedly less volatile under rational inattention as compared to the perfect information RBC model. This is expected given the low information flow agents in the economy allocate to the aggregate environment and the consequent dampening. Another effect of rational inattention in an otherwise standard RBC setting is that the delay in the response of aggregate variables leads to stronger autocorrelations and cross-correlations.

Variable	SD (%)	x(-4)	x(-3)	x(-2)	x(-1)	x	x(+1)	x(+2)	x(+3)	x(+4)
		Cross Correlation of Output with : (Full Information)								
C	0.62	0.50	0.59	0.69	0.78	0.85	0.54	0.28	0.07	-0.09
I	11.61	0.12	0.28	0.48	0.72	0.99	0.78	0.59	0.41	0.27
L	1.72	0.10	0.26	0.46	0.71	0.99	0.78	0.60	0.43	0.29
Y	2.39	0.21	0.36	0.55	0.76	1.00	0.76	0.55	0.36	0.20
		Cross Correlation of Output with : (Rational Inattention)								
Variable	**SD (%)**	x(-4)	x(-3)	x(-2)	x(-1)	x	x(+1)	x(+2)	x(+3)	x(+4)
C	0.6	0.59	0.68	0.75	0.82	0.87	0.66	0.45	0.25	0.06
I	7.84	0.26	0.44	0.63	0.82	0.98	0.87	0.74	0.60	0.46
L	0.8	0.48	0.66	0.82	0.92	0.93	0.80	0.66	0.52	0.37
Y	1.75	0.37	0.53	0.70	0.86	1.00	0.86	0.70	0.53	0.37

Table 3. Business Cycle Statistics - Perfect Information vs Rational Inattention

5.2.1 Comparing Two Different TFP Volatility Regimes: Great Moderation as a Case Study

Figure 4 plots the impulse responses of aggregate variables to an innovation in aggregate TFP under different TFP-volatility regimes and different information structures. The "high volatility" impulse responses correspond to an economy with aggregate TFP calibrated to the US data prior to 1984. The "low volatility" impulse responses correspond to an economy with TFP calibrated to the post-1984 period. Following the evidence of Fernald (2009), I assume that TFP innovations are 34% less volatile post 1984. As the economy moves from high to low aggregate TFP volatility, the impulse responses of output and hours experience a bigger change under rational inattention as compared to full information. As the economy is hit by less volatile aggregate TFP shocks, firms optimally choose to reallocate their attention towards tracking idiosyncratic TFP, and therefore respond less to innovations in aggregate TFP. This is the mechanism that leads to the *volatility amplification* effect.

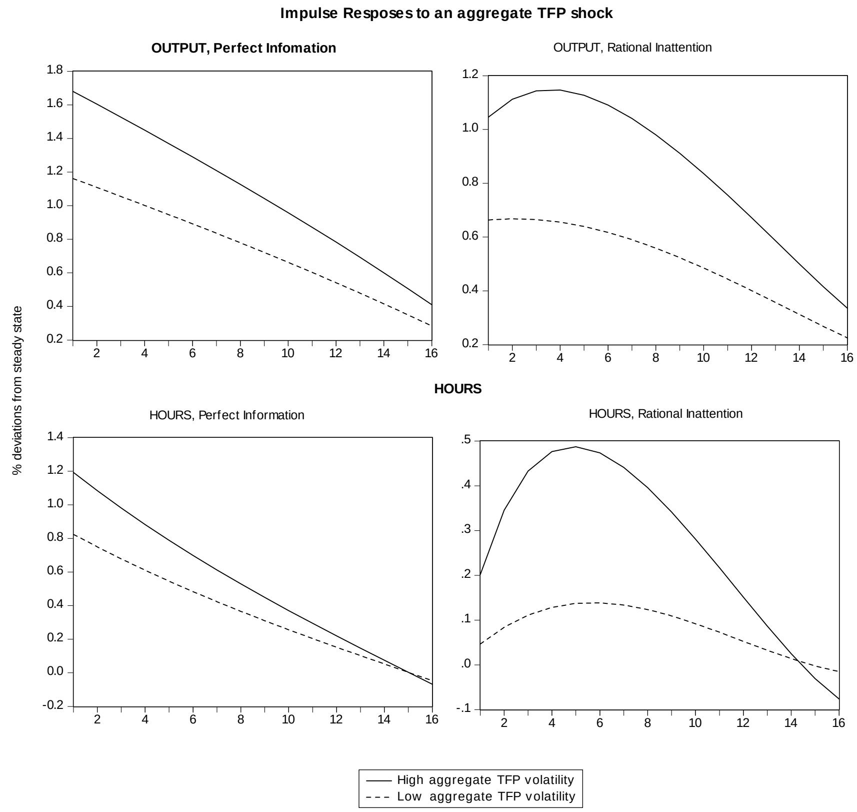

Figure 4. Impulse Responses to an aggregate TFP shock across different TFP volatility regime and information structures

The magnitude of this amplification effect, which is the main result of this paper, is summarized in Table 4. I simulate the models 200 times, with each simulation consisting of 300 periods. I then HP filter the simulated data and compute the volatility of output, hours, consumption and investment. For the model under rational inattention, a 34% decline in the standard deviation of the innovation to aggregate TFP leads to a 46% decline in the volatility of aggregate output, a 72% decline in the volatility of hours, a 33% decline in the volatility of consumption and a 50% decline in the volatility of investment. Under perfect information, when aggregated, the model collapses to a standard RBC model with decreasing returns to scale. In that case a 34% decline in aggregate TFP volatility leads to only 34% decline in the volatility of all macroeconomic variables. Hence, the model under rational inattention differs from the full information model in two ways. First, it amplifies changes in the volatility of aggregate TFP. Second, the response to changes in the volatility of TFP is different across aggregate variables. It is stronger for hours and weaker for consumption. The lack of volatility amplification for consumption is because for simplicity households are assumed to have infinite information processing capacity, i.e. perfect information about the state of the economy. The reason why volatility of hours responds more than that of output under rational inattention but not under perfect information can be explained as follows. Under perfect information both labor and output depend on the *true* state of technology (aggregate TFP). Under rational inattention hours depend on the *perceived* state of technology ($E[a_t|s^t]$), whereas output is determined by the true state of technology as well as hours employed in production according to the production function. Changes in the volatility of aggregate TFP lead to bigger changes in the volatility of the perceived state, as the latter is a function of attention allocation. Because output is a function of these two states (a_t and $E[a_t|s^t]$), in percentage terms its volatility will change by more than the change in TFP volatility and by less than the change in hours volatility. See Appendix C for the proof.

Table 4. Great Moderation: Data versus RBC and Rational Inattention (RI)
(percent standard deviations)

Series	**Output**	**Hours**	**Consumption**	**Investment**
Data (1961 - 2006)	1.55	1.78	0.78	4.56
Data (1961 - 1983)	1.90	2.01	0.92	5.41
Data (1983 - 2006)	0.94	1.44	0.56	3.15
Data (late/early)	**0.49**	**0.72**	**0.61**	**0.58**
Rational Inattention (pre 1984)	1.75	0.80	0.60	7.84
Rational Inattention (post 1984)	0.95	0.33	0.40	3.92
RI (late/early)	**0.54**	**0.28**	**0.67**	**0.50**
RBC (pre 1984)	2.39	1.72	0.62	11.61
RBC (post 1984)	1.58	1.14	0.41	7.65
RBC (late/early)	**0.66**	**0.66**	**0.66**	**0.66**
$\sigma_\varepsilon(pre1984) = 0.0092, \sigma_\varepsilon(post1984) = 0.006, \frac{\sigma_\varepsilon(post1984)}{\sigma_\varepsilon(pre1984)} = 0.66$				

6 Shutting Down the Idiosyncratic Channel: Rational Inattention versus Attention Allocation

In this section I explore the extent to which allowing for idiosyncratic volatility matters for aggregate dynamics. There are two dimensions of rational inattention that are important for this paper. First, firms have imperfect information about the state vector due to their limited ability to process information. Second, the presence of the idiosyncratic shocks forces the firms to allocate attention to tracking the idiosyncratic state, at the cost of less information being allocated to the aggregate environment. Changes in the volatility of idiosyncratic and/or aggregate shocks do not affect the total precision of firms' signals, but do affect the way precision is allocated across signals. The direction in which the relative volatility of the shocks changes determines the direction of attention reallocation. In the case where there is no idiosyncratic volatility to compete for attention, all information processing capacity will be allocated to improving the precision of signals regarding the aggregate state. In this case a change in the volatility of aggregate shocks does not change the amount of information flow that goes to tracking the true state of the economy. In such an environment there is no volatility amplification effect.

6.1 Rational Inattention Problem for the Firm

To illustrate the importance of idiosyncratic volatility to my results, I examine an alternative model in which firms face only aggregate shocks, but are still subject to imperfect information in the form of a capacity constraint on per period information flow. My setting is the standard RBC model with an information processing constraint placed on the side of the representative firm.

$$\min E\left[\sum_{t=0}^{\infty}\beta^t(\frac{\hat{\pi}_{33}}{2}(\hat{k}_t^*-\hat{k}_t^F)^2+\frac{\hat{\pi}_{44}}{2}(\hat{l}_t^*-\hat{l}_t^F)^2+\hat{\pi}_{34}(\hat{k}_t^*-\hat{k}_t^F)(\hat{l}_t^*-\hat{l}_t^F))\right] \tag{51}$$

subject to

$$\hat{l}_t^F = \frac{1}{1-\alpha-\delta}(a_t-(1-\alpha)\hat{w}_t-\alpha\hat{r}_t) \tag{52}$$

$$\hat{k}_t^F = \frac{1}{1-\alpha-\delta}(a_t-\delta_t\hat{w}_t-(1-\delta)\hat{r}_t) \tag{53}$$

$$\hat{l}_t^* = E\left[\hat{l}_t^F|s_i^t\right] \tag{54}$$

$$\hat{k}_t^* = E\left[\hat{k}_t^F|s_i^t\right] \tag{55}$$

$$I(\{w_t,r_t,a\};\{s_{it}\})\leq\kappa \tag{56}$$

If we remove the most important shock (idiosyncratic shock) and hold κ constant, firms will have enough information flow to track the aggregate shock almost perfectly and the results under rational inattention and perfect information will be indistinguishable. There will be no delay or dampening in the responses of hours, output and investment to an innovation in aggregate TFP, and there will be no volatility amplification. This is only due to the fact that firms have an abundance of information processing ability on their hands.

To make the exercise interesting, suppose instead that agents are endowed with much less information processing capacity than in the benchmark model. In particular, suppose κ equals 0.23 bits, which is the amount of information flow per period allocated to aggregate shocks in the benchmark model. In this case rational inattention will lead to dampened and delayed responses in aggregate outcomes to the aggregate technology shock, but there will be no volatility amplification. This is due to the fact that changes in underlying shock volatility do not lead to changes in the information flow allocated to that shock (since it is the only shock). To make this point clear, I set $\kappa = 0.23$ in the imperfect information model with only aggregate shocks and compare its volatility amplification effects (if any) with the benchmark and the RBC models. Table 5 shows that even when the model under Rational Inattention with only aggregate shocks is calibrated to yield less volatility than the RBC model, it still maintains a linear relationship between the volatility of the aggregate shock and the volatility of aggregate outcomes. That is, a 34% decline in the volatility of the aggregate technology shock leads to 34% decline in the volatility of aggregate variables just as in the standard perfect information RBC model.

Table 5. Rational inattention (RI) without the attention allocation problem versus standard RBC model.

(percent standard deviations)

Series	**Output**	**Hours**	**Consumption**	**Investment**
Data (1961 - 2006)	1.55	1.78	0.78	4.56
Data (1961 - 1983)	1.90	2.01	0.92	5.41
Data (1983 - 2006)	0.94	1.44	0.56	3.15
Data (late/early)	**0.49**	**0.72**	**0.61**	**0.58**
Rational Inattention (pre 1984)	1.75	0.80	0.60	7.84
Rational Inattention (post 1984)	1.15	0.53	0.39	5.13
RI (late/early)	**0.66**	**0.66**	**0.66**	**0.66**
RBC (pre 1984)	2.39	1.72	0.62	11.61
RBC (post 1984)	1.58	1.14	0.41	7.65
RBC (late/early)	**0.66**	**0.66**	**0.66**	**0.66**
$\sigma_\varepsilon(pre-1984)=.0092, \sigma_\varepsilon(post-1984)=.006, \frac{\sigma_\varepsilon(post-1984)}{\sigma_\varepsilon(pre-1984)}=.66$				

7 Can Changes in the Volatility of the Idiosyncratic Environment Cause Changes in the Macroeconomic Environment ?

In this section I ask whether changes in the idiosyncratic shock process alone can generate changes in the dynamics of macroeconomic aggregates. In the following numerical exercise I examine how an economy under rational inattention responds to an increase in the volatility of idiosyncratic shocks. The "low volatility" impulse responses correspond to an economy with idiosyncratic TFP calibrated to US data prior to 1984. The "high volatility" impulse responses correspond to an economy with idiosyncratic TFP being hypothetically 25% more volatile. Everything else is kept unchanged.

Figure 5 plots the impulse responses of output and hours to an innovation in aggregate TFP when the economy moves from a low-volatility to a high-volatility idiosyncratic environment under rational inattention and perfect information. Under perfect information, the response of variables to an innovation in aggregate TFP is the same under high or low idiosyncratic volatility. That is, under perfect information, the nature of the idiosyncratic environment plays no role for aggregate dynamics. On the other hand, under rational inattention, the volatility of the idiosyncratic environment matters for the aggregate dynamics. The more volatile the idiosyncratic shock, the more dampened the response of aggregate variables to an innovation in aggregate TFP, as shown in the right hand column in Figure 6.

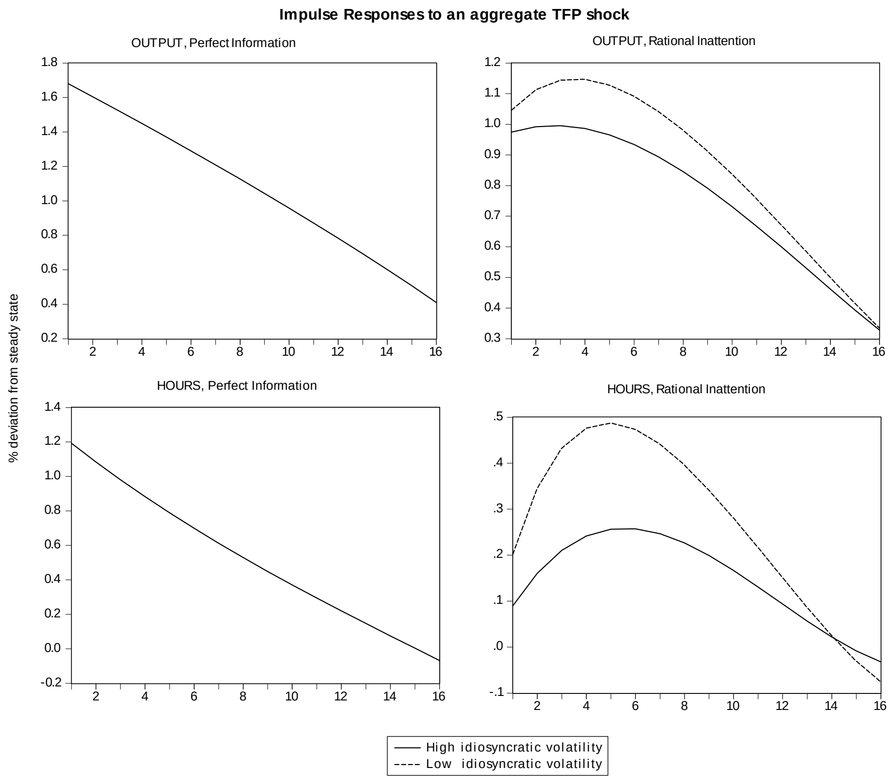

Figure 5. Impulse response of output and hours to an innovation in aggregate TFP across different idiosyncratic volatility regimes

Table 6 shows the magnitude of the decline in aggregate volatility due to an hypothetical 25% increase in the standard deviation of the innovations in the idiosyncratic TFP. The perfect information case as expected is not affected by changes in the idiosyncratic environment. However, the rational inattention case offers a role for the idiosyncratic environment in aggregate dynamics. Changes in idiosyncratic volatility change the allocation of attention, which affects the equilibrium behavior of agents in the economy. In other words, the transmission mechanism of aggregate shocks in the economy is a function in part of the stochastic properties governing idiosyncratic shocks. Keeping all other benchmark parameters unchanged, an *increase* of 25 % in the standard deviation of idiosyncratic shocks leads to a 11% *decline* in volatility of aggregate output and a 36% decline in that of aggregate hours.

Table 6. 25% increase in idiosyncratic TFP volatility no change in aggregate TFP volatility (percent standard deviations)				
Series	**Output**	**Hours**	**Consumption**	**Investment**
RI^{low}	1.75	0.80	0.60	7.84
RI^{high}	1.56	0.51	0.6	6.69
$\mathbf{RI}^{high}/\mathbf{RI}^{low}$	**0.89**	**0.64**	**1.00**	**0.85**
RBC^{low}	2.39	1.72	0.62	11.61
RBC^{high}	2.39	1.72	0.62	11.61
$\mathbf{RBC}^{high}/\mathbf{RBC}^{low}$	1.00	1.00	1.00	1.00
$\sigma_u(high) = .2242, \sigma_u(low) = .1746, \sigma_\varepsilon(high) = \sigma_\varepsilon(low)$				

Reconciling a contemporaneous increase in idiosyncratic volatility and a decrease in macroeconomic volatility is of particular importance when looking at another established fact during the Great Moderation episode, which is the increased household-level consumption and income volatility (Gottschalk and Moffitt (2002),Comin, Groshen, and Rabin (2006), Hyslop (2001)). Increased household level volatility in the mid 1980s in the face of a decline in macroeconomic activity during the same period has stirred considerable research.

Augmenting the benchmark model with rational inattention in the side of the consumers as well as firms, could potentially reconcile the contemporaneous increase in household level volatility and the decline in macroeconomic volatility. I will pursue this extension of my model in my future research.

8 Comparative statics: changing structural parameters

8.1 Changing the upper bound on information flow

Rational inattention theory relies on the assumption that agents in the economy are limited in information processing ability and face an upper bound on the maximum information flow that can be processed. As the maximum amount of information that agents in the economy can process (κ) increases, the model under rational inattention approaches that of full information. Figure 6 shows the relative response of aggregate hours and output to an innovation in aggregate TFP under rational inattention as compared to perfect information[16]. This ratio approaches 1 as the capacity to process information (κ) increases.

[16] I report results in this section using the labor-only model of section 4.2 to save computational time. Preliminary experiments suggest that qualitative results carry over to the benchmark model with capital. Future drafts will report results for the benchmark model.

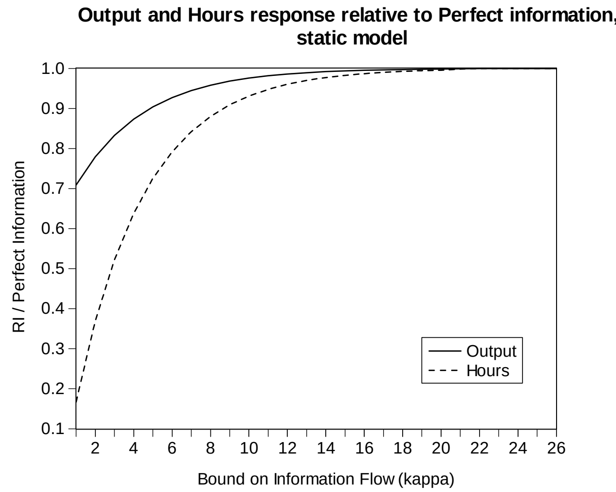

Figure 6. Output and Hours time-0 response relative to perfect information, static model.

Figure 7 displays computed volatility elasticities for output and hours for the labor-only model. The elasticity of the volatility of hours with respect to the volatility of the aggregate TFP shock is higher than that for output. Both elasticities approach the perfect information outcome of unit-elasticity if the upper bound in the information processing capacity is higher than 6 bits per quarter. Both figures 6 and 7 illustrate the fact that as the upper bound on per period information flow increases, the model under rational inattention approaches the perfect information model.

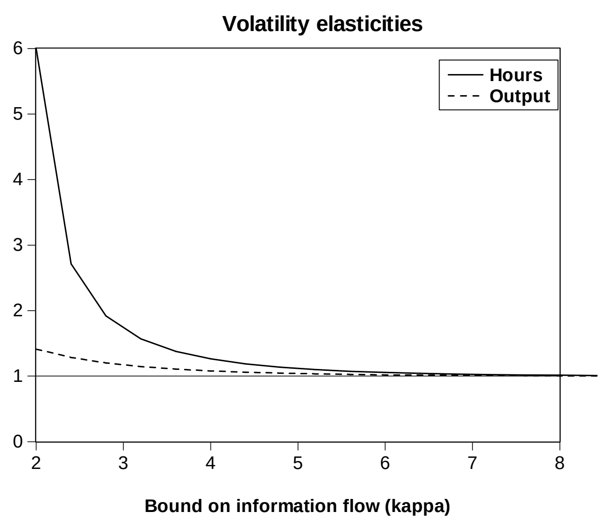

Figure 7. Volatility elasticities for Hours and Output, labor-only model

8.2 Different household preferences

Here I explore the implications of the form of household preferences for volatility amplification. I compare the results for the benchmark separable preferences versus the preferences assumed in Greenwood-Hercowitz-Hoffman (GHH, 1988).[17] The specification of preferences determines the dynamics on the labor supply side of the economy and hence affects the feedback mechanism between imperfect information on the side of the firms and the household sector. The GHH preference function is as follows:

$$U(C_t, L_t) = \frac{(C_t - \psi L_t^{\nu})^{1-\sigma}}{1-\sigma}, \ \psi > 0, \ \nu > 1$$

Whereas the preferences in the benchmark model are:

$$U(C_t, L_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma} - \theta \frac{L_t^{1+\psi}}{1+\psi}$$

The main difference between these two types of preferences is the equilibrium labor supply. Under GHH preferences, labor supply is independent of consumption, due to the absence of wealth effects. Both preference specifications lead to a volatility amplification effect, but of different magnitude. In the numerical experiments, I calibrate the two different models such that they yield the same steady state equilibrium. I find that the amplification is smaller in magnitude for GHH preferences. The absence of wealth effects leads to less reallocation of attention in response to a change in the volatility of aggregate shocks. The intuition is the following: when the economy faces a decline in the volatility of aggregate shocks, this will lead firms in all cases to reallocate attention away from the aggregate environment, which will be reflected in the weights they put on various shocks in their demand for inputs. Such changes in the input demand by firms will have to be matched by changes in the input supply of households. Under GHH preferences labor supply responds differently to changes in labor demand than under the benchmark preference specification. In particular, the change in labor supply is accomplished only through a change in the wage rate rather than consumption. For preference specifications with wealth effects and hence a negative covariance between consumption and labor supply, a larger change in the wage rate will be required to match a given change in the demand for labor by firms. This leads to bigger volatility amplification for preference specifications which allow for wealth effects.

9 Conclusion

In a standard RBC model there is an almost linear relationship between the volatility of aggregate TFP shocks and the volatility of aggregate variables such as output, employment and investment.[18] This paper shows that endogenizing the information set in an otherwise standard RBC model breaks this linear relationship. Following the literature on rational inattention, agents in this economy are assumed to be constrained in their ability to process information and face the decision of how to

[17] I consider Cobb-Douglas preferences as well. Results show that amplification is similar for separable and Cobb-Douglas preferences.

[18] This relationship is exactly linear up to a first order approximation and very close to linear for higher order appoximations.

allocate this limited information flow across many state variables of interest. The trade-off they face in terms of allocating limited attention across aggregate and idiosyncratic states is the key aspect of the model that leads to a non-linear relationship between the volatility of aggregate TFP and macroeconomic variables. The observed 34% decline in TFP volatility from the pre-1984 to the post-1984 period can generate a 46% decline in output volatility when agents rationally reallocate attention away from aggregate shocks and towards idiosyncratic shocks.

Hence, rational inattention with attention allocation implies that equi-proportional changes in the volatility of aggregate shocks are not necessary to generate a given magnitude of change in the volatility of macroeconomic variables. One of the key variables that determines the extent of this non-linear relationship between TFP volatility and output volatility is the relative volatility of aggregate versus idiosyncratic shocks. This variable determines how much attention is allocated to each state variable, with more information flow being directed towards the nosier variable. Hence, a relatively more noisy idiosyncratic environment would lead to more attention being allocated towards idiosyncratic states at the cost of less information being allocated to aggregate shocks. The contribution of this paper is to bring forth the importance of endogenous information sets as well as the interaction between the aggregate and idiosyncratic environment in determining macroeconomic volatility. There are several extension of this model that I intend to work in the future.

First, this model can be extended to allow for rational inattention on the side of consumers as well as firms. This would be particularly interesting since this model could reconcile two established facts regarding the 1984-2006 period, that of increasing household level earnings volatility and declining macroeconomic volatility (Gottschalk and Moffitt (2002), Comin, Groshen, and Rabin (2006), Hyslop (2001)). As shown in Section 7 of the paper, the attention allocation mechanism can lead to a contemporaneous increase in idiosyncratic volatility and a decline in aggregate volatility.

A second extension of this model would be to allow for monetary shocks as another aggregate shock in the economy. The reason for this is to address the observed decline in inflation volatility that the U.S. has experienced during 1984-2006. This would be complementary to the Maćkowiak and Wiederholt (2009b) DSGE model of rational inattention where they allow for technology and monetary policy shocks.

Third, this model can be extended to allow for a time variation in the volatility of the structural innovations. This would have implications for the time-variation in the share of information allocated across shocks.

References

[1] Adam, K. (2007). "Optimal Monetary Policy with Imperfect Common Knowledge," Journal of Monetary Economics 54(2), 276–301.

[2] Ahmed, S., A. Levin and B.Wilson (2004). "Recent U.S. Macroeconomic Stability: Good Luck, Good Policies, or Good Practices?" Review of Economics and Statistics 86(3), 824-32.

[3] Arias, A., G. Hansen, and L. Ohanian (2006). "Why Have Business Cycle Fluctuations Become Less Volatile?" NBER Working Paper #12079.

[4] Basu, Susanto, John Fernald, and Miles Kimball (2006). "Are Technology Improvements Contractionary?" American Economic Review, vol. 96(5), 1418-1448.

[5] Bullard, James B. and Singh, Aarti (2007). "Learning and the Great Moderation," Working Paper 2007-027a, Federal Reserve Bank of St. Louis.

[6] Campbell, J. R., and J. D. M. Fisher (2004). "Idiosyncratic Risk and Aggregate Employment Dynamics," Review of Economic Dynamics vol. 7(2), 331-353.

[7] Clarida, R., Galí, J. & Gertler, M. (2000). "Monetary policy rules and macroeconomic stability: Evidence and some theory," Quarterly Journal of Economics 115, 147–180.

[8] Comin, Diego, Erica L. Groshen, and Bess Rabin (2006). "Turbulent Firms, Turbulent Wages?" National Bureau of Economic Research Working Paper #12032.

[9] Comin, D., and T. Philippon (2005). "The Rise in Firm-Level Volatility: Causes and Consequences," In M. Gertler and K. Rogoff, eds. NBER Macroeconomics Annual Volume 20, 167-202. Cambridge, MA: MTT Press.

[10] Cooper, R., Haltiwanger, J., Willis, J. (2007). " Search frictions: Matching aggregate and establishment observations," Journal of Monetary Economics 54, 56–78.

[11] Davis, Steven J., John C. Haltiwanger, Ron Jarmin and Javier Miranda (2006). "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," NBER Macroeconomics Annual.

[12] Dynan, Karen, Douglas W. Elmendorf, and Daniel E. Sichel (2006). "Financial Innovation and the Great Moderation What Do the Household Data Say?", conference on "Financial Innovations and the Real Economy".

[13] Fernald, John (2009). "A Quarterly, Utilization-Corrected Series on Total Factor Productivity," mimeo, Federal Reserve Bank of San Francisco.

[14] Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" American Economic Review, 98(1), 394–425

[15] Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Hoffman (1988). "Investment, Capacity Utilization, and the Real Business Cycle," American Economic Review, 78(3) June: 402-17.

[16] Hopenhayn, Hugo A. (1992). "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," Econometrica, 60(5) 1127–50.

[17] Jaimovich, Nir and Rebelo, Sergio T.(2006). "Can News about the Future Drive the Business Cycle?" CEPR Discussion Paper No. 5877.

[18] Hyslop, Dean R. (2001). “Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings,” American Economic Review, 755-777.

[19] Justiniano, A., and G. Primiceri (2006). "The Time-Varying Volatility of Macroeconomic Fluctuations," NBER Working Paper #12022.

[20] Luo, Yulei (2008). "Consumption Dynamics under Information Processing Constraints," Review of Economic Dynamics, 11, 366-385.

[21] Luo, Yulei and Young, Eric R. (2009) "Rational Inattention and Aggregate Fluctuations," The B.E. Journal of Macroeconomics: Vol. 9 (1) (Contributions), Article 14.

[22] M. S. Pinsker (1964). Information and information stability of random variables and processes, Holden Day, San Francisco.

[23] Maćkowiak ,B., Wiederholt, M. (2009a). "Optimal sticky prices under rational inattention," American Economic Review 99, 769–803

[24] Maćkowiak, B., Wiederholt, M. (2009b). "Business cycle dynamics under rational inattention," Discussion paper, European Central Bank and Northwestern University.

[25] McConnell, M. M. & Perez-Quiros, G. (2000), "Output Fluctuations in the united states: What has changed since the early 1980's?", American Economic Review 90(5), 1464–1476.

[26] Moffitt, Robert A. and Peter Gottschalk (2002). “Trends in the Transitory Variance of Earning in the United States,” The Economic Journal, C68-C73

[27] Paciello, Luigi (2008). "The Response of Prices to Technology and Monetary Policy Shocks under Rational Inattention," Discussion paper, Northwestern University.

[28] Restuccia, Diego and Rogerson, Richard (2004). "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," Society for Economic Dynamics, working paper, no. 69.

[29] Sims, C., and T. Zha (2006). "Were There Regime Switches in U.S. Monetary Policy?" American Economic Review 96(1), 54-81.

[30] Sims, Christopher A. (1998). "Stickiness." Carnegie-Rochester Conference Series on Public Policy, 49, 317–56.

[31] Sims, Christopher A. (2003). "Implications of Rational Inattention." Journal of Monetary Economics, 50(3), 665–90.

[32] Sims, Christopher A. (2006). "Rational Inattention: Beyond the Linear-Quadratic Case," American Economic Review Papers and Proceedings, 96(2): 158-163.

[33] Steven J. Davis, John Haltiwanger, Ron Jarmin and Javier Miranda (2006). "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms", NBER working paper 12354

[34] Stock, J., and M. Watson (2003). “Has the Business Cycle Changed and Why?” NBER Macroeconomics Annual 2002 17, 159-218.

Appendices

A DATA

Data on macroeconomic aggregates are taken from Federal Reserve Economic Data (FRED) dataset and Bureau of Labor Statistics (BLS) The data series include seasonally adjusted, quarterly, billions of chained 2000$, real gross national product, real personal consumption expenditures of durable, non-durable goods and services, real private fixed investment, hours and employment.

B Non-stochastic steady state

In the deterministic steady state there are no technology shocks : $a_{it} = a_t = 0$. Given that technology is the only source of heterogeneity in the model, in this case all firms are exactly the same.

From the household first order conditions we have:

$$\bar{C}^{-\gamma} w = \theta \bar{L} \tag{57}$$

$$1 = \beta(1 + r - d) \tag{58}$$

For the representative firm (due to lack of heterogeneity in the deterministic steady state) we have:

$$w = \delta \bar{K}^{\alpha} \bar{L}^{-\delta - 1} \tag{59}$$

$$r = \alpha \bar{K}^{\alpha - 1} \bar{L}^{\delta} \tag{60}$$

From the aggregate resource constraint and the production function we have:

$$\bar{C} = \bar{Y} + d\bar{K} \tag{61}$$

$$\bar{Y} = \bar{K}^{\alpha} \bar{L}^{\delta} \tag{62}$$

There are 6 equations and 6 unknowns, so I can solve for $\left\{\bar{Y}, \bar{C}, \bar{K}, \bar{L}, w, \bar{r}\right\}$.

C Why volatility amplification is stronger for aggregate hours of work than aggregate output

Suppose $Y_t = g(z_t, L_t)$, where $z_t = e^{a_t}$ and $g(.)$ is any production function. After log-linearizing output around $z_t = 1, L_t = \bar{L}$ we have:

$$\hat{Y}_t = \frac{g_z(1,\bar{L})}{\bar{Y}} a_t + \frac{g_L(1,\bar{L})}{\bar{Y}} \hat{L}_t$$

Under rational inattention $\hat{L}_t = f(\frac{\sigma_u}{\sigma_\varepsilon}) a_t$. Assume for simplicity that $a_t = \varepsilon_t$. Then we have: $\hat{Y}_t = \left(\frac{g_z(1,\bar{L})}{\bar{Y}} + \frac{g_L(1,\bar{L})}{\bar{Y}} f(\frac{\sigma_u}{\sigma_\varepsilon})\right) \varepsilon_t$. The volatilities of labor and output are

$$Var(L_t) = f(\frac{\sigma_u}{\sigma_\varepsilon})^2 \sigma_\varepsilon^2$$

and

$$Var(\hat{Y}_t) = \left(\frac{g_z(1,\bar{L})}{\bar{Y}} + \frac{g_L(1,\bar{L})}{\bar{Y}} f(\frac{\sigma_u}{\sigma_\varepsilon})\right)^2 \sigma_\varepsilon^2$$

The elasticities of $Var(L_t)$ and $Var(\hat{Y}_t)$ with respect to σ_ε^2 are :

$$\epsilon_{\sigma_\varepsilon^2}^{var(L)} = 1 + \frac{2 f_{\sigma_\varepsilon^2}(.) \sigma_\varepsilon^2}{f(.)}$$

and

$$\epsilon_{\sigma_\varepsilon^2}^{var(Y)} = 1 + \frac{2 f_{\sigma_\varepsilon^2}(.) \sigma_\varepsilon^2}{\frac{g_z(1,\bar{L})}{g_L(1,\bar{L})} + f(.)}$$

Given that $\frac{g_z(1,\bar{L})}{g_L(1,\bar{L})}$ is always positive,

$$\epsilon_{\sigma_\varepsilon^2}^{var(Y)} < \epsilon_{\sigma_\varepsilon^2}^{var(L)}$$

D Derivation of the information flow constraint

In this subsection I will derive the information rate for one and two-dimensional discrete parameter Gaussian processes using frequency-domain methods.

D.1 Information rate of discrete parameter one-dimensional Gaussian processes

Let $X = \{x(t)\}, Y = \{y(t)\}$ be one-dimensional, real-valued, discrete parameter , wide-sense stationary and stationarily correlated processes. The information rate between these two processes can be written as follows

$$I_{X,Y} = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \log(1 - |r_{XY}(\omega)|^2) d\omega$$

where

$$|r_{XY}(\omega)|^2 = \{ ^{\frac{|f_{XY}(\omega)|^2}{f_{XX}(\omega) f_{YY}(\omega)}, f_{XY}(\omega) \neq 0}_{0, f_{XY}(\omega)=0}$$

where, $f_{XX}(\omega)$ and $f_{YY}(\omega)$ are spectral densities of process X and Y respectively, and $f_{XY}(\omega)$ is the cross-spectral density. $|r_{XY}(\omega)|^2$ is also called the coherence between the processes at frequency ω, which is the frequency-domain analog of the correlation coefficient.

As an example of this, assume that X and Y can be expressed as infinite-order moving average: $X = \sum_{l=0}^{\infty} d_l \varepsilon_{t-l} = D(L)\varepsilon_t$ and $Y = \sum_{l=0}^{\infty} m_l^L \varepsilon_{t-l} + \sum_{l=0}^{\infty} n_l^L \eta_{t-l}^L(L)\varepsilon_t = M^L(L)\varepsilon_t + N^L(L)\eta_t^L$, where $D(L), M^L(L), N^L(L)$ are infinite lag polynomials and $\{\varepsilon_t\}, \{\eta_t^L\}$ are Gaussian mutually independent white noise processes with σ_ε^2 and unit variance respectively and independent of each other. Spectral density functions for X_1 and Y_1 are:

$$f_{XX}(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} D(e^{-i\omega}) D(e^{i\omega})$$

$$f_{YY}(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} M^L(e^{-i\omega}) M^L(e^{i\omega}) + \frac{1}{2\pi} N^L(e^{-i\omega}) N^L(e^{i\omega})$$

and the cross-spectral density is

$$f_{XY}(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} D(e^{-i\omega}) M^L(e^{i\omega})$$

where $D(e^{-i\omega}) = d_o + d_1 e^{-i\omega} + d_2 e^{-2i\omega} + ... d_T e^{-Ti\omega} + ..,$ $D(e^{i\omega}) = d_o + d_1 e^{i\omega} + d_2 e^{2i\omega} + ... d_T e^{Ti\omega} + ..,$ $M^L(e^{-i\omega}) = m_o^L + m_1^L e^{-i\omega} + m_2^L e^{-2i\omega} + ... m_T^L e^{-Ti\omega} + ..,$ $M^L(e^{i\omega}) = m_o^L + m_1^L e^{i\omega} + m_2^L e^{2i\omega} + ... m_T^L e^{Ti\omega} + ..$ and $N^L(e^{-i\omega}) = n_o^L + n_1^L e^{-i\omega} + n_2^L e^{-2i\omega} + ... n_T^L e^{-Ti\omega} + ..,$ $N^L(e^{i\omega}) = n_o^L + n_1^L e^{i\omega} + n_2^L e^{2i\omega} + ... n_T^L e^{Ti\omega} + ...$ Using the spectral and cross-spectral densities, the information rate between these two one-dimensional processes becomes:

$$I_{X,Y} = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \log\left(\frac{1}{1 + \frac{\sigma_\varepsilon^2 M^L(e^{-i\omega}) M^L(e^{i\omega})}{N^L(e^{-i\omega}) N^L(e^{i\omega})}}\right) d\omega$$

where $\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega}) M^L(e^{i\omega})}{N^L(e^{-i\omega}) N^L(e^{i\omega})}$ is also defined as the signal-to-noise ratio. Hence, one can express the information rate between two moving average Gaussian processes in terms of their moving average coefficients. This information flow constraint will be used in the dynamic version of the model with labor only as the input choice to be made by the firms.

D.2 Information rate of discrete parameter multi-dimensional Gaussian processes[19]

The multidimensional case of the problem applies to the benchmark model in the paper, where the firms' optimal input choices are those of capital and labor.

Let $X = \{x_1(t), x_2(t), ...x_n(t)\}, Y = \{y_1(t), y_2(t), ...y_m(t)\}$ be n and m-dimensional, real-valued,

[19] Derivations in this section follow the book "Information and information stability of random variables and processes" by M. S. Pinsker (1964)

discrete parameter , wide-sense stationary and stationarily correlated processes respectively. The information rate between these two processes can be written as follows:

$$I_{X,Y} = -\frac{1}{4\pi}\int_{-\pi}^{\pi} \log \frac{\det A_{\tilde{X}\tilde{Y}}(\omega)}{\det A_{\tilde{X}}(\omega)\det A_{\tilde{Y}}(\omega)} d\omega$$

where $\det A_X(\omega) = \det ||f_{x_i x_j}(\omega)||_{i,j=1,.....,n}$, $\det A_Y(\omega) = \det ||f_{y_i y_j}(\omega)||_{i,j=1,.....,m}$, $\det A_{XY}(\omega) = \det ||f_{x_i y_j}(\omega)||_{i,j=1,.....,n+m}$ and $\det A_{\tilde{X}}(\omega)$ is a non-vanishing principal minor of highest order $'r'$ of the determinant $\det A_X(\omega)$, $\det A_{\tilde{Y}}(\omega)$ is a non-vanishing principal minor of highest order $'s'$ of the determinant $\det A_Y(\omega)$, and $\det A_{\tilde{X}\tilde{Y}}(\omega)$ is the principal minor of order $'r+s'$ of the determinant $\det A_{XY}(\omega)$ which contains $\det A_{\tilde{X}}(\omega)$ and $\det A_{\tilde{Y}}(\omega)$. $f_{12}(\omega)$ refers to the cross-spectrum between variable '1' and '2'.

The model in this paper requires the computation of the information rate between two-dimensional Gaussian processes. The information flow relevant in the model is the information flow between the full information profit maximizing decisions of capital and labor, and the actual decisions under limited information. In turn, this can be interpreted as the information rate between the variable the firms are trying to track (the profit maximizing decisions) and the signals they get regarding the profit maximizing decisions, which are the actual decisions.

We have $I(\{l_{it}^F\},\{k_{it}^F\};\{l_{it}^*\},\{k_{it}^*\}) = I(\{l_t^{FA}\},\{k_t^{FA}\};\{l_t^{*A}\},\{k_t^{*A}\}) + I(\{l_{it}^{FI}\},\{k_{it}^{FI}\};\{l_{it}^{*I}\},\{k_{it}^{*I}\})$, where subscript F stands for full information optimal decisions and subscript $*$ stands for actual decisions for capital and labor, and where A stands for aggregate components while I stands for the idiosyncratic components. The equality above comes from the fact that common and idiosyncratic components of the firm-level productivity shock are independent from each other. Hence, I can separate the aggregate from the idiosyncratic component in each decision rule[20]. In order to compute the information flow, I use the moving average representation of decision rules for capital and labor derived under full and incomplete information. The following derivation involves the information flow pertaining to the aggregate component of the decision rules.

$l_t^{FA} = D(L)\varepsilon_t$, $k_t^{FA} = E(L)\varepsilon_t$, $l_{it}^{*A} = M^L(L)\varepsilon_t + N^L(L)\eta_{it}^L$, $k_{it}^{*A} = M^K(L)\varepsilon_t + N^K(L)\eta_{it}^K$, where $\varepsilon_t \sim WN(0,\sigma_\varepsilon^2), \eta_{it}^L$ and $\eta_{it}^K \sim WN(0,1)$, where $\{\varepsilon_t\},\{\eta_{it}^L\}$ and $\{\varepsilon_t\}$, $\{\eta_{it}^K\}$ are pairwise independent from each other but $\{\eta_{it}^L\},\{\eta_{it}^K\}$ do not need to be independent. This setting applies to an environment where there is a single agent (the firm's decision maker) that chooses the optimal pair of labor and capital inputs. The objective of the firm is to track the full information profit-maximizing levels of labor and capital using an optimal set of signals. Since there is only one decision maker within the firm that jointly chooses labor and capital inputs, it is reasonable to assume that information processing will lead to optimal signals being correlated. This paper allows for this possibility, which expands the set of choice variables for the firm when they solve their attention allocation problem. Firms now will choose not only the extent of the noise in each signal but also their correlation across signals.

After calculating the spectral and cross-spectral densities as well as using the definition for

[20] This same procedure is followed in Maćkowiak and Wiederholt (2009a)

information flow for multi-dimensional Gaussian processes I obtain:

$$I(\{l_t^{FA}\}, \{k_t^{FA}\}; \{l_t^{*A}\}, \{k_t^{*A}\}) =$$

$$-\frac{1}{4\pi}\int_{-\pi}^{\pi} \log \frac{1}{1+\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega})M^L(e^{i\omega})}{(1-\chi^2)N^L(e^{-i\omega})N^L(e^{i\omega})}+\frac{\sigma_\varepsilon^2 M^{K-i\omega})M^K(e^{i\omega})}{(1-\chi^2)N^K(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{-i\omega})M^K(e^{i\omega})}{N^L(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{i\omega})M^K(e-^{i\omega})}{N^L(e^{i\omega})N^K(e-^{i\omega})}} d\omega$$

where $\chi = E(\eta_{it}^L \eta_{it}^K)$.

By looking at the profit-maximizing decision rules for each firm , the idiosyncratic component for both labor and capital input decisions is the same, namely the idiosyncratic TFP component. In this case the firm chooses to receive only one signal whose noise will be a choice variable.

$l_{it}^{FI} = k_t^{FI} = A_2(L)u_t$, $l_{it}^{*I} = k_{it}^{*I} = S(L)u_{it} + T(L)\psi_{it}$, where $u_{it} \sim WN(0, \sigma_u^2), \psi_{it} \sim WN(0, 1)$,

$$I(\{l_{it}^{FI}\}, \{k_{it}^{FI}\}; \{l_{it}^{*I}\}, \{k_{it}^{*I}\} = -\frac{1}{4\pi}\int_{-\pi}^{\pi} \log \frac{1}{1 + \frac{\sigma_u^2 S(e^{-i\omega})S(e^{i\omega})}{T(e^{-i\omega})T(e^{i\omega})}} d\omega$$

E Algorithm

The algorithm used here to solve the model is similar to Paciello (2008).

Step 1:

Under both types of information structures, I solve the model by log-linearizing around the deterministic steady-state. It is well-known that under full-information log-linearization, eliminates second-moment effects. However, under incomplete information with information processing constraints, there are first-order effects of the volatility of underlying shocks, even though the model is log-linearized.

Full Information

Under full-information the following equations must hold in equilibrium:

$$\psi \hat{L}_t + \gamma \hat{C}_t = \hat{w}_t$$

$$\hat{C}_t = E(\hat{C}_{t+1} - \frac{\hat{r}_{t+1}}{\gamma})$$

$$\hat{Y}_t = \frac{\bar{C}}{\bar{Y}}\hat{C}_t + \frac{\bar{K}}{\bar{Y}}(\hat{K}_{t+1} - (1-d)\hat{K}_t)$$

$$\hat{l}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t)$$

$$\hat{k}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - \delta\hat{w}_t - (1-\delta)\hat{r}_t)$$

$$a_{it} = \rho_I a_{it-1} + u_{it},\ u_{it} \sim WN(0, \sigma_u^2)$$

$$a_t = \rho_A a_t + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

The first two equations come from household problem, the third one the resource constraint, the third is from the resource constraint, the fourth and the fifth equations are optimal labor and capital decisions taken by firms under full-information, and the last two equations are the assumed processes for the common and idiosyncratic components of firm-level TFP. Given the assumption of decreasing returns to scale one can determine optimal hours of work and capital, unlike the case of constant returns to scale, where only the capital-to-labor ratio can be pinned down. Part of step 1 involves making a guess for the deviation of capital and labor decisions under rational inattention from the profit-maximizing decisions (under full information)[21]. The guess takes the following form: $guess^L = l^*_{it} - l^F_{it}$ and $guess^K = k^*_{it} - k^F_{it}$

Using the guess I compute the implied dynamics for the model for the aggregate variables. The set of equations that must hold in equilibrium for the aggregate dynamics under rational inattention are the following:

$$\psi \hat{L}_t + \gamma \hat{C}_t = \hat{w}_t$$

$$\hat{C}_t = E(\hat{C}_{t+1} - \frac{\hat{r}_{t+1}}{\gamma})$$

$$\hat{Y}_t = \frac{\bar{C}}{\bar{Y}}\hat{C}_t + \frac{\bar{K}}{\bar{Y}}(\hat{K}_{t+1} - (1-d)\hat{K}_t)$$

$$y_t = a_t + \delta l_t + \alpha k_t$$

$$a_t = \rho_A a_t + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

Obtaining the average wage and rental rate I can compute the profit-maximizing decision rules for capital and labor, which are used in solving the attention allocation problem:

$$l_t^{FA} = \frac{1}{1-\alpha-\delta}(a_t - (1-\alpha)w_t - \alpha r_t)$$

and

$$k_t^{FA} = \frac{1}{1-\alpha-\delta}(a_t - \delta w_t - (1-\delta) r_t)$$

One can express all variables as moving averages. For instance, $a_t = A_1(L)\varepsilon_t$, $w_t = W(L)\varepsilon_t$,

[21] This step is similar to formulating a guess regarding the actual labor and capital decisions under rational inattention.

$r_t = R(L)\varepsilon_t$. Substituting these moving average representations into l_t^{FA} and k_t^{FA} I obtain: $l_t^{FA} = D(L)\varepsilon_t$, $k_t^{FA} = E(L)\varepsilon_t$, where $D(L) = \frac{1}{1-\alpha-\delta}(A_1(L) - (1-\alpha)W(L) - \alpha R(L))\varepsilon_t$ and $E(L) = \frac{1}{1-\alpha-\delta}(a_t - \delta W(L) - (1-\delta)R(L))\varepsilon_t$. The idiosyncratic part of the profit-maximizing decision rules is simply $l_{it}^{FI} = k_{it}^{FI} = \frac{1}{1-\alpha-\delta}a_{it} = \frac{1}{1-\alpha-\delta}A_2(L)u_{it}$, where $A_2(L)u_{it}$ is a moving average representation of the idiosyncratic component of the firm-level TFP shock.

Step 2. Having obtained the profit -maximizing decision rules for capital and labor I can now solve the attention allocation problem that firms face. Each firm minimizes the losses it incurs due to incomplete information, subject to an information processing constraint.

$$Loss = \frac{1}{2}E[\pi_{33}(k_{it} - k_{it}^F)^2 + 2\pi_{34}(k_{it} - k_{it}^F)(l_{it} - l_{it}^F) + \pi_{44}(l_{it} - l_{it}^F)^2] =$$

$$\frac{1}{2}E[\pi_{33}(k_{it}^A - k_{it}^{FA})^2 + \pi_{33}(k_{it}^I - k_{it}^{FI})^2 + \pi_{44}(l_{it}^A - l_{it}^{FA})^2 + \pi_{44}(l_{it}^I - l_{it}^{FI})^2$$

$$+2\pi_{34}(k_{it}^A - k_{it}^{FA})(l_{it}^A - l_{it}^{FA}) + 2\pi_{34}(k_{it}^I - k_{it}^{FI})(l_{it}^I - l_{it}^{FI})] =$$

$$\frac{1}{2}E[\pi_{33}(k_{it}^A - k_{it}^{FA})^2 + \pi_{44}(l_{it}^A - l_{it}^{FA})^2 + 2\pi_{34}(k_{it}^A - k_{it}^{FA})(l_{it}^A - l_{it}^{FA})]$$

$$+\frac{1}{2}E[\pi_{33}(k_{it}^I - k_{it}^{FI})^2 + \pi_{44}(l_{it}^I - l_{it}^{FI})^2 + 2\pi_{34}(k_{it}^I - k_{it}^{FI})(l_{it}^I - l_{it}^{FI})]$$

where

$$\begin{aligned} l_t^{FA} &= D(L)\varepsilon_t \\ k_t^{FA} &= E(L)\varepsilon_t \\ l_{it}^{*A} &= M^L(L)\varepsilon_t + N^L(L)\eta_{it}^L \\ k_{it}^{*A} &= M^K(L)\varepsilon_t + N^K(L)\eta_{it}^K \end{aligned} \tag{63}$$

where $\varepsilon_t \sim WN(0,\sigma_\varepsilon^2), \eta_{it}^L$ and $\eta_{it}^K \sim WN(0,1)$,where $\{\varepsilon_t\}, \{\eta_{it}^L\}$ and$\{\varepsilon_t\}$, $\{\eta_{it}^K\}$ are pairwise independent and $E(\eta_{it}^L\eta_{it}^K) = \chi$.

$$\begin{aligned} l_{it}^{FI} &= k_t^{FI} = A_2(L)u_t \\ l_{it}^{*I} &= k_{it}^{*I} = S(L)u_{it} + T(L)\psi_{it} \end{aligned} \tag{64}$$

where $u_{it} \sim WN(0,\sigma_u^2)$,and $\psi_{it} \sim WN(0,1)$. Lag polynomials $D(L)$ and $E(L)$ come from step 1 given the initial guess whereas the moving average coefficients on the actual decisions are what the firms choose.

Information flow can also be expressed as the sum of information flow between idiosyncratic variables and information flow between aggregate variables.

$$I(\{l_{it}^F\},\{k_{it}^F\};\{l_{it}^*\},\{k_{it}^*\}) = I(\{l_t^{FA}\},\{k_t^{FA}\};\{l_t^{*A}\},\{k_t^{*A}\}) + I(\{l_{it}^{FI}\},\{k_{it}^{FI}\};\{l_{it}^{*I}\},\{k_{it}^{*I}\})$$

$$= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\frac{1}{1+\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega})M^L(e^{i\omega})}{(1-\chi^2)N^L(e^{-i\omega})N^L(e^{i\omega})}+\frac{\sigma_\varepsilon^2 M^{K-i\omega)}M^K(e^{i\omega})}{(1-\chi^2)N^K(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{-i\omega})M^K(e^{i\omega})}{N^L(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{i\omega})M^K(e{-i\omega})}{N^L(e^{i\omega})N^K(e{-i\omega})}}d\omega$$
$$-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\frac{1}{1+\frac{\sigma_u^2 S(e^{-i\omega})S(e^{i\omega})}{T(e^{-i\omega})T(e^{i\omega})}}d\omega$$

The attention allocation problem becomes:

$$\max_{\{m^K,m^L,n^K,n^L,s,t\}}\frac{1}{2}(\frac{1}{1-\alpha-\delta})^2\{\sigma_\varepsilon^2\pi_{33}\sum_{l=0}^T(m_l^K-e_l)^2+\pi_{33}\sum_{l=0}^T(n_l^K)^2+$$

$$\sigma_\varepsilon^2\pi_{44}\sum_{l=0}^T(m_l^L-d_l)^2+\pi_{44}\sum_{l=0}^T(n_l^L)^2+2\pi_{34}\sigma_\varepsilon^2\sum_{l=0}^T(m_l^K-e_l)(m_l^L-d_l)+$$

$$2\pi_{34}\chi\sum_{l=0}^T n_l^K n_l^L$$

$$\sigma_u^2(\pi_{44}+\pi_{33}+2\pi_{34})\sum_{l=0}^T(s_l-a_{2l})^2+$$

$$(\pi_{44}+\pi_{33}+2\pi_{34})\sum_{l=0}^T(t_l)^2\}$$

subject to

$$-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\frac{1}{1+\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega})M^L(e^{i\omega})}{(1-\chi^2)N^L(e^{-i\omega})N^L(e^{i\omega})}+\frac{\sigma_\varepsilon^2 M^{K-i\omega)}M^K(e^{i\omega})}{(1-\chi^2)N^K(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{-i\omega})M^K(e^{i\omega})}{N^L(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{i\omega})M^K(e{-i\omega})}{N^L(e^{i\omega})N^K(e{-i\omega})}}d\omega$$

$$-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\frac{1}{1+\frac{\sigma_u^2 S(e^{-i\omega})S(e^{i\omega})}{T(e^{-i\omega})T(e^{i\omega})}}d\omega\leq\kappa$$

where $\{m^K,m^L,n^K,n^L,s,t\}$ are the lag polynomial coefficients in equations (63) and (64).

As previously derived, the information flow is a function of moving average coefficients, which also appear in the loss function. As an example, consider the choice of m^L, n^L :

$$(\frac{1}{1-\alpha-\delta})^2\sigma_\varepsilon^2\pi_{44}(m_l^L-d_l)+\pi_{34}\sigma_\varepsilon^2(m_l^K-e_l)=$$

$$-\frac{\lambda}{4\pi}\int_{\pi}^{-\pi}\frac{\partial\left(\log\frac{1}{1+\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega})M^L(e^{i\omega})}{(1-\chi^2)N^L(e^{-i\omega})N^L(e^{i\omega})}+\frac{\sigma_\varepsilon^2 M^{K-i\omega)}M^K(e^{i\omega})}{(1-\chi^2)N^K(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{-i\omega})M^K(e^{i\omega})}{N^L(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{i\omega})M^K(e{-i\omega})}{N^L(e^{i\omega})N^K(e{-i\omega})}}\right)}{\partial m_l^L}$$

and

$$(\frac{1}{1-\alpha-\delta})^2(\pi_{44}+2\pi_{34}\chi n_l^K)n_l^L =$$

$$-\frac{\lambda}{4\pi}\int_{-\pi}^{\pi}\frac{\partial\left(\log\frac{1}{1+\frac{\sigma_\varepsilon^2 M^L(e^{-i\omega})M^L(e^{i\omega})}{(1-\chi^2)N^L(e^{-i\omega})N^L(e^{i\omega})}+\frac{\sigma_\varepsilon^2 M^{K-i\omega)}M^K(e^{i\omega})}{(1-\chi^2)N^K(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{-i\omega})M^K(e^{i\omega})}{N^L(e^{-i\omega})N^K(e^{i\omega})}-\frac{\sigma_\varepsilon^2\chi}{1-\chi^2}\frac{M^L(e^{i\omega})M^K(e^{-i\omega})}{N^L(e^{i\omega})N^K(e^{-i\omega})}}\right)}{\partial n_l^L}$$

where λ is the shadow price of information. The complete solution of the attention allocation stage consists of 6T+1 equations and 6T+1 unknowns, which are solved numerically. Once this stage is solved I obtain$\{l_{it}^*\}\{k_{it}^*\}$, which are the actual decisions under rational inattention. As a next step I compute the difference between these decision rules and profit maximizing decision rules. If $l_{it}^* - l_{it}^F \neq guess^L$ and $k_{it}^* - k_{it}^F \neq guess^K$ I update the guess by the following rule:

$$guess_{new}^L = \phi guess^L + (1-\phi)(l_{it}^* - l_{it}^F)$$

and

$$guess_{new}^K = \phi guess^K + (1-\phi)(k_{it}^* - k_{it}^F)$$

F Perfect Information Case

In this section I compute the equilibrium dynamics of the full-information version of the model in which firms know the entire history of state variables, including their period t realization. Under full information the model collapses to a standard RBC model with DRTS technology in the production function. Hence, the perfect information solution is not only important in comparing the two different information structures but also because it nests a well known benchmark, that of a standard RBC model.

The household part of the economy is the same as in the benchmark model. Given that there are no adjustment costs to the firm of changing the number of workers or capital, their problem is static.

The firm's problem is:

$$\max_{l_{it},k_{it}}\left\{e^{a_t}e^{a_{it}}k_{it}^{\alpha}l_{it}^{\delta} - w_t l_{it} - r_t k_{it}\right\} \tag{65}$$

The implied first order conditions are:

$$w_t = \delta e^{a_t}e^{a_{it}}k_{it}^{\alpha}l_{it}^{\delta-1} \tag{66}$$

$$r_t = \alpha e^{a_t}e^{a_{it}}k_{it}^{\alpha-1}l_{it}^{\delta} \tag{67}$$

Which implies :

$$\frac{w_t}{r_t} = \left(\frac{\delta}{\alpha}\right)\frac{k_{it}}{l_{it}} \tag{68}$$

All firms have the same capital-to-labor ratio. The DRTS assumption allows me to pin down firm-specific levels of labor and capital demand:

$$l_{it} = \left[\frac{\delta e^{a_t} e^{a_{it}} (\frac{\alpha w_t}{\delta r_t})^{\alpha}}{w_t}\right]^{\frac{1}{1-\alpha-\delta}} \tag{69}$$

$$k_{it} = l_{it}(\frac{w_t}{r_t}\frac{\alpha}{\delta}) \tag{70}$$

The market clearing conditions are $K_t = \int k_{it}di,\ L_t = \int l_{it}di,\ Y_t = \int y_{it}di, \int a_{it}di = 0$.

The resource constraint is:

$$C_t + K_{t+1} - (1-d)K_t = Y_t \tag{71}$$

Log-linearized version of the Perfect Information Model

Given that the imperfect information model will be solved in a Linear Quadratic Gaussian framework, I need the log-linearized FOC of the perfect information case to make a consistent comparison as well to build a quadratic loss function. The log-linearization is done around the non-stochastic steady state (see Appendix A).

The log-linearized set of first order conditions for the household and firms are:

$$\psi \hat{L}_t + \gamma \hat{C}_t = \hat{w}_t \tag{72}$$

$$\hat{C}_t = E\left(\hat{C}_{t+1} - \frac{\hat{r}_{t+1}}{\gamma}\right) \tag{73}$$

$$\hat{w}_t = a_t + a_{it} + \alpha \hat{k}_{it}^F + (\delta - 1)\hat{l}_{it}^F \tag{74}$$

$$\hat{r}_t = a_t + a_{it} + (\alpha - 1)\hat{k}_{it}^F + \delta \hat{l}_{it}^F \tag{75}$$

$$\hat{k}_{it}^F - \hat{l}_{it}^F = \hat{w}_t - \hat{r}_t \tag{76}$$

$$\hat{y}_{it}^F = a_t + a_{it} + \alpha \hat{k}_{it}^F + \delta \hat{l}_{it}^F \tag{77}$$

$$\hat{l}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha \hat{r}_t) \tag{78}$$

$$\hat{k}_{it}^F = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - \delta \hat{W}_t - (1-\delta)\hat{r}_t) \tag{79}$$

$$\hat{Y}_t = \frac{\bar{C}}{\bar{Y}}\hat{C}_t + \frac{\bar{K}}{\bar{Y}}\left(\hat{K}_{t+1} - (1-d)\hat{K}_t\right) \tag{80}$$

Aggregate Equilibrium Conditions under Perfect Information

By aggregating the firm-specific first order conditions we obtain 6 equations, three of which are equations (72), (73) and (80), and 6 unknowns $\{L_t, K_{t+1}, C_t, Y_t, w_t, r_t\}$:

$$\hat{Y}_t = a_t + \alpha\hat{K}_t + \delta\hat{L}_t \tag{81}$$

$$\hat{L}_t = \frac{1}{1-\alpha-\delta}\left(a_t - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t\right) \tag{82}$$

$$\hat{K}_t = \frac{1}{1-\alpha-\delta}(a_t - \delta\hat{W}_t - (1-\delta)\hat{r}_t) \tag{83}$$