ABSTRACT

Title of dissertation: ESSAYS ON MACROECONOMIC VOLATILITY AND MONETARY ECONOMICS

Jeta Menkulasi, Doctor of Philosophy, 2010

Dissertation directed by: Professor Boragan Aruoba
Professor John Haltiwanger

Department of Economics

My dissertation consists of two independent essays on macroeconomic volatility and monetary economics respectively. The first essay explores the implications of imperfect information on macroeconomic volatility. It 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.

The second essay examines the redistributive effects of monetary policy using a dynamic general equilibrium model with heterogenous agents. I study the long-run effects of inflation on output, consumption and welfare, as well as the distribution of wealth in the economy. Unlike in representative agent models, heterogeneity can potentially allow for beneficial effects of inflation. Increases in the growth rate of money supply can reduce wealth dispersion, increasing output and welfare.

ESSAYS ON MACROECONOMIC VOLATILITY AND MONETARY ECONOMICS

by

Jeta Menkulasi

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy
2010

Advisory Committee:

Professor Boragan Aruoba (co-chair)
Professor John Haltiwanger (co-chair)
Professor Curt Grimm
Professor Anton Korinek
Professor John Shea

Dedication

To my mother, Kozeta.

Acknowledgements

This dissertation is the end result of a great support from a number of people. Foremost, I would like to thank my advisors, Professor Boragan Aruoba and Professor John Haltiwanger, the co-chairs to this dissertation committee, for their guidance and extensive time devoted to discussing my research.

I am highly indebted to Prof. Boragan Aruoba for his support over the years. He has been constantly available to discuss any questions arising in my research in great details. He has provided invaluable help with conceptual as well as technical obstacles I have encountered. Most important, he has encouraged me when I have been most critical of my work. I can never thank him enough for his patience and his persistence.

I am very grateful to Professor John Haltiwanger for his insights and help with data central to my research. I also thank Professor John Shea for taking his time to read in great detail my drafts and provide excellent technical comments.

My work has benefited extensively from the seminars in the Department of Economics. I would like to thank Prof. Anton Korinek, Prof. Carlos Vegh and Prof. Allan Drazen for their challenging questions, which have improved the quality of my research. I am particularly grateful to Professor Enrique Mendoza for taking his time to discuss my work.

I would also like to thank Professor Curt Grimm for accepting to serve as an external advisor to my committee.

This dissertation comes at the cost of being away from my family for years. My family members, my father, Gazmen, my mother, Kozeta and my brother, Fatmir have been a constant moral support throughout this period. I am lucky to be the daughter of Kozeta Menkulasi and the granddaughter of Gjystina Dishnica, who have been my source of strength whenever I doubted myself.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Rational Inattention and Changes in Macroeconomic Volatility

1.1 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 chapter 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

[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.

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 chapter 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

[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.

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 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 chapter 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 chapter 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,

[4]See Maćkowiak and Wiederholt (2009a)

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 firm-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 1.1).[5] Using indirect inference, I estimate a similar (9%) decline in

[5] Figure 1.1 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.

the volatility of idiosyncratic TFP, which combined with the 34% decline in aggregate TFP volatility, 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. This chapter 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 reallocation 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 chapter. 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

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

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 chapter 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 chapter discovers a new outcome of rational inattention, which is a volatility amplification effect. One main contribution of this chapter 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 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 chapter overlaps with their paper 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.

Overall the contribution of this chapter in the literature is twofold. First, it is the first paper to expose a volatility-amplification result in rational inattention models

[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.

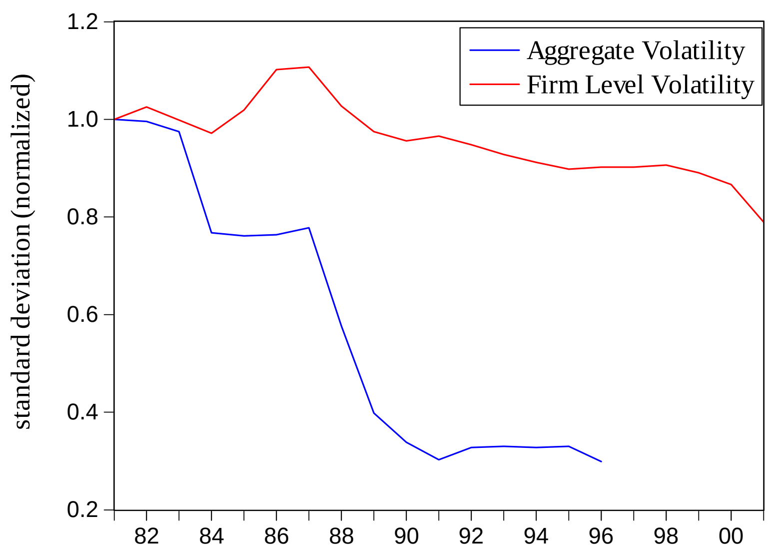
with attention allocation. Second, it offers a new application of the rational inattention theory.

This chapter is organized as follows: section 1.2 introduces the tools from information theory that are applied in my rational inattention setting. Section 1.3 introduces the benchmark model. In section 1.4, I study a simple version of the model that has an analytical solution to illustrate the main mechanism in this chapter. Section 1.5 presents the calibration procedure and the numerical results for the benchmark model. In section 1.6, 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 1.7 examines whether changes in the volatility of the idiosyncratic environment alone can lead to changes in aggregate volatility. Section 1.8 includes the sensitivity analysis. Section 1.9 analyses the implications of a model with endogenously determined upper bound on the ability to process information. Section 1.10 concludes.

1.2 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,

Figure 1.1: Comparison of Patterns of Firm and Aggregate Volatility using Employment Growth Rates from LBD



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

1.3 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.

1.3.1 Firms

This part of the model is similar to Restuccia and Rogerson (2004) as well as Bartelsman, Haltiwanger and Scarpetta (2009) 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. Obtaining a non-degenerate distribution of firm size is important in supporting a distribution of the idiosyncratic productivity in equilibrium, and hence, exploring the role of the idiosyncratic environment.

The production technology each firm faces is

$$y_{it} = e^{a_t} e^{a_{it}} k_{it}^{\alpha} l_{it}^{\delta} \quad , \alpha + \delta < 1 \tag{1.1}$$

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

$$a_t = \rho_A a_{t-1} + \varepsilon_t \tag{1.2}$$

$$a_{it} = \rho_I a_{it-1} + u_t \tag{1.3}$$

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

Profits in each period are

$$\pi(k_{it,}l_{it,}w_{t,}r_t) = e^{a_t}e^{a_{it}}k_{it}^{\alpha}l_{it}^{\delta} - w_tl_{it} - r_tk_{it} \tag{1.4}$$

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.[8] 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

$$\max_{\{s_{it}\}\in S}E\left[\sum_{t=0}^{\infty}\tilde{\beta}_t\Pi(k_{it,}^*l_{it,}^*w_{t,}r_t, a_{t,}a_{it})\right] \tag{1.5}$$

subject to

$$I(\{w_{t,}r_t, a_{t,}a_{it}\};\{s_{it}\}) \leq \kappa \tag{1.6}$$

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

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. 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 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 chapter. 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).[9]

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.[10]

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.

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

[10] See Maćkowiak and Wiederholt (2009a)

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*

$$\hat{l}_{it}^F = \frac{1}{1-\alpha-\delta}\left[a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t\right] \tag{1.7}$$

$$\hat{k}_{it}^F = \frac{1}{1-\alpha-\delta}\left[a_t + a_{it} - \delta\hat{w}_t - (1-\delta)\hat{r}_t\right] \tag{1.8}$$

and aggregate labor and capital follow

$$\hat{L}_t = \frac{1}{1-\alpha-\delta}\left[a_t - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t\right] \tag{1.9}$$

$$\hat{K}_t = \frac{1}{1-\alpha-\delta}\left[a_t - \delta\hat{w}_t - (1-\delta)\hat{r}_t\right] \tag{1.10}$$

Proof. See Appendix G. ■

It is important to emphasize that under perfect information, the aggregate economy looks exactly like the representative agent RBC model with decreasing returns to scale (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.

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

$$\tilde{\pi}(a_{t,} a_{it,} \hat{k}_{it,} \hat{l}_{it,} \hat{w}_t, \hat{r}_t) \simeq \hat{\pi}(0,0,0,0,0,0,0) + \hat{\pi}_1 a_t + \hat{\pi}_2 a_{it} + \hat{\pi}_3 \hat{k}_{it} + \hat{\pi}_4 \hat{l}_{it} + \hat{\pi}_5 \hat{w}_t + \hat{\pi}_6 \hat{r}_t$$
$$+ \tfrac{\hat{\pi}_{11}}{2} a_t^2 + \tfrac{\hat{\pi}_{22}}{2} a_{it}^2 + \tfrac{\hat{\pi}_{33}}{2} \hat{k}_{it}^2 + \tfrac{\hat{\pi}_{44}}{2} \hat{l}_{it}^2 + \tfrac{\hat{\pi}_{55}}{2} \hat{w}_t + \tfrac{\hat{\pi}_{66}}{2} \hat{r}_t$$
$$+ \hat{\pi}_{12} a_t a_{it} + \hat{\pi}_{13} a_t \hat{k}_{it} + \hat{\pi}_{14} a_t \hat{l}_{it} + \hat{\pi}_{15} a_t \hat{w}_t + \hat{\pi}_{16} a_t \hat{r}_t$$
$$+ \hat{\pi}_{23} a_{it} \hat{k}_{it} + \hat{\pi}_{24} a_{it} \hat{l}_{it} + \hat{\pi}_{25} a_{it} \hat{w}_t + \hat{\pi}_{26} a_{it} \hat{r}_t$$
$$+ \hat{\pi}_{34} \hat{k}_{it} \hat{l}_{it} + \hat{\pi}_{35} \hat{k}_{it} \hat{w}_t + \hat{\pi}_{36} \hat{k}_{it} \hat{r}_t + \hat{\pi}_{45} \hat{l}_{it} \hat{w}_t + \hat{\pi}_{46} \hat{l}_{it} \hat{r}_t + \hat{\pi}_{56} \hat{w}_t \hat{r}_t$$

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

$$\hat{l}_{it}^* = \phi_a^L E[a_t|s_i^t] + \phi_I^L E[a_{it}|s_i^t] + \phi_w^L E[w_t|s_i^t] + \phi_r^L E[r_t|s_i^t] \tag{1.11}$$

$$\hat{k}_{it}^* = \phi_a^K E[a_t|s_i^t] + \phi_I^K E[a_{it}|s_i^t] + \phi_w^K E[w_t|s_i^t] + \phi_r^K E[r_t|s_i^t] \tag{1.12}$$

where $\{k_{it}^*, l_{it}^*\}$ stand for optimal capital and labor input under rational inattention.[11]

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

$$\hat{l}_{it}^F = \phi_a^L a_t + \phi_I^L a_{it} + \phi_w^L w_t + \phi_r^L r_t \tag{1.13}$$

$$\hat{k}_{it}^F = \phi_a^K a_t + \phi_I^K a_{it} + \phi_w^K w_t + \phi_r^K r_t \tag{1.14}$$

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 $\hat{k}_{it}^* =$

[11]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).

$E\left[\hat{k}_{it}^F | s_i^t\right]$. 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 (1.13), (1.14) 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

$$\min_{\{s_{it}\}} E\left\{\sum_{t=0}^{\infty} \beta^t \left[\frac{\hat{\pi}_{33}}{2}(\hat{k}_{it}^* - \hat{k}_{it}^F)^2 + \frac{\hat{\pi}_{44}}{2}(\hat{l}_{it}^* - \hat{l}_{it}^F)^2 + \hat{\pi}_{34}(\hat{k}_{it}^* - \hat{k}_{it}^F)(\hat{l}_{it}^* - \hat{l}_{it}^F)\right]\right\} \quad (1.15)$$

subject to

$$\hat{l}_{it}^F = \tfrac{1}{1-\alpha-\delta}(a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha\hat{r}_t) \tag{1.16}$$

$$\hat{k}_{it}^F = \tfrac{1}{1-\alpha-\delta}(a_t + a_{it} - \delta_t\hat{w}_t - (1-\delta)\hat{r}_t) \tag{1.17}$$

$$\hat{l}_{it}^* = E\left[\hat{l}_{it}^F | s_i^t\right] \tag{1.18}$$

$$\hat{k}_{it}^* = E\left[\hat{k}_{it}^F | s_i^t\right] \tag{1.19}$$

$$I(\{w_{t,}r_t, a_{t,}a_{it}\}; \{s_{it}\}) \leq \kappa \tag{1.20}$$

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 (1.20). 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 E.

1.3.2 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

$$C_t + K_{t+1} = w_t L_t + (1 + r_t - d)K_t + \Pi_t \tag{1.21}$$

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

$$\lim_{T\to\infty} E_0[\pi_{t=0}^{T}(1 + r_{t+1})^{-1}]K_{T+1} = 0 \tag{1.22}$$

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

$$C_t^{-\gamma} w_t = \theta L_t^{\psi} \tag{1.23}$$

$$C_t^{-\gamma} = \beta E_t[C_{t+1}^{-\gamma}(1 + r_{t+1} - d)] \tag{1.24}$$

1.3.3 Equilibrium

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

$$C_t^{-\gamma} w_t = \theta L_t^{\psi} \tag{1.25}$$

$$C_t^{-\gamma} = \beta E_t[C_{t+1}^{-\gamma}(1 + r_{t+1} - d)] \tag{1.26}$$

the resource constraint:

$$C_t + K_{t+1} - (1-d)K_t = Y_t \tag{1.27}$$

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

$$L^s(w_t, r_t, a_t) = \int_i L^d(s_i^t)di \tag{1.28}$$

$$K^s(w_t, r_t, a_t) = \int_i K^d(s_i^t)di \tag{1.29}$$

market clearing condition

$$Y_t = \int y_{it}di$$

and the aggregate and idiosyncratic components of firm-level TFP, which are assumed to follow independent AR(1) processes.

$$a_{it} = \rho_I a_{it-1} + u_{it} \tag{1.30}$$

$$a_t = \rho_A a_t + \varepsilon_t \tag{1.31}$$

$$\textstyle\int a_{it}di = 0 \tag{1.32}$$

1.4 Special Case: No Capital and White Noise Disturbances

In order to illustrate the main mechanism in the model here I solve a special case of the incomplete information model where disturbances follow a white noise process and capital is fixed. 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 yields an analytic solution, which clarifies the main mechanism in the chapter.

1.4.1 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

$$\hat{L}_t^F = \frac{1-\gamma}{1+\psi-\delta+\delta\gamma} a_t \tag{1.33}$$

$$\hat{w}_t^F = \frac{\psi+\gamma}{1+\psi-\delta+\delta\gamma} a_t \tag{1.34}$$

$$\hat{C}_t^F = \frac{1+\psi}{1+\psi-\delta+\delta\gamma} a_t \tag{1.35}$$

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 I will show analytically how macroeconomic dynamics under rational inattention do depend on the idiosyncratic environment and how this leads to a volatility amplification effect.

1.4.2 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

$$\min_{\{s_{it}\}} E\left[\sum_{t=0}^{\infty} \beta^t \ \frac{\pi_{33}}{2} \ (\hat{l}_{it}^* - \hat{l}_{it}^F)^2\right] \tag{1.36}$$

subject to

$$\hat{l}_{it}^{F} = \frac{1}{1-\delta}\left(a_t + a_{it} - \hat{w}_t\right) \tag{1.37}$$

$$\hat{l}_{it}^{*} = \frac{1}{1-\delta}(E\left[a_t|s_i^t\right] + E\left[a_{it}|s_i^t\right] - E\left[\hat{w}_t|s_i^t\right]) \tag{1.38}$$

$$I\left(\left\{\hat{w}_t, a_{t,} a_{it}\right\}; \left\{s_{it}\right\}\right) \leq \kappa \tag{1.39}$$

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 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".[12] Hence, we have

$$s_{1it} = a_t + u_{it} \tag{1.40}$$

$$s_{2it} = a_{it} + \varepsilon_{it} \tag{1.41}$$

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

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

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

$$
\begin{aligned}
E(a_t|s_{1it}) &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_u^2}s_{1it} \\
E(a_{it}|s_{2it}) &= \frac{\sigma_{a_i}^2}{\sigma_{a_i}^2+\sigma_\varepsilon^2}s_{2it}
\end{aligned}
$$

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

$$
\min_{\frac{\sigma_a^2}{\sigma_u^2},\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}} \left\{ \left(\frac{1-\varphi}{1-\delta}\right)^2 \sigma_a^2 \left[\left(\frac{1}{\frac{\sigma_a^2}{\sigma_u^2}+1}\right)^2 + \frac{1}{\frac{\sigma_a^2}{\sigma_u^2}}\right] + \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 \left[\left(\frac{1}{\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}+1}\right)^2 + \frac{1}{\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}}\right]\right\} \tag{1.42}
$$

subject to

$$
\frac{1}{2}\log_2\left(1+\frac{\sigma_a^2}{\sigma_u^2}\right)+\frac{1}{2}\log_2\left(1+\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}\right) \le \kappa \tag{1.43}
$$

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

Optimal signal-to-noise ratios for each signal are

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

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

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

$$L_t = \frac{\varphi - \gamma}{\psi + \gamma\delta} a_t \tag{1.44}$$

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 (1.44)

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

Using the assumptions on the signals and the derived information flow constraint, the *interior* solution to the attention allocation problem is as follows[13]

$$\hat{L}_t = \hat{L}_t^F \left(1 - \frac{1}{1-\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}2^{-\kappa}}\right) \tag{1.45}$$

$$\hat{C}_t = \hat{C}_t^F \left(1 - \frac{\delta}{1+\psi}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}2^{-\kappa}}\right) \tag{1.46}$$

$$\hat{w}_t = \hat{w}_t^F \left(1 - \frac{\psi+\gamma\delta}{\psi+\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}2^{-\kappa}}\right) \tag{1.47}$$

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 (1.33), (1.34) and (1.35).

The solution under rational inattention, equations (1.45), (1.46) and (1.47), 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

[13]See appendix for details on these derivations.

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

$$\epsilon_{\sigma_a^2}^{var(Y)} = \frac{1}{1 - \frac{\delta}{1+\psi}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}}2^{-\kappa}} > 1 \tag{1.48}$$

$$\epsilon_{\sigma_a^2}^{var(L)} = \frac{1}{1 - \frac{1}{1-\gamma}\sqrt{\frac{\sigma_{ai}^2}{\sigma_a^2}}2^{-\kappa}} > 1 \tag{1.49}$$

These elasticities are the main concern of this chapter. 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

$$\kappa^A = \frac{1}{2}\log_2\left[\frac{1}{1-\delta+\psi+\gamma\delta}\left(\frac{\sigma_a}{\sigma_{a_i}}(1-\delta)(1-\gamma)+\ 2^{-\kappa}(\psi+\gamma\delta)\right)\right]+\frac{1}{2}\kappa \qquad (1.50)$$

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

$$\kappa^I = \kappa - \kappa^A \qquad (1.51)$$

Equation (1.50) 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.

1.5 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.

1.5.1 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.[14] Parameters governing the persistence and standard deviation of the aggregate TFP shock are obtained using the quarterly series on TFP computed by Fernald (2007). I fit equation (1.31) 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.

Idiosyncratic TFP process

I use the evidence on firm-level data compiled by 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

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

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. There is little evidence on firm-level shocks but Foster, Haltiwanger and Syverson (2008) provide direct persistence estimates of plant-level TFP shocks, which are around 0.80. Cooper and Haltiwanger (2006), also using plant-level data estimate the persistence parameter of the idiosyncratic shock to be around 0.89. In this chapter, in the absence of the relevant firm-level data required to compute the idiosyncratic TFP, I conduct an indirect inference exercise. I match the model's predictions for firm-level employment dynamics with moments from firm-level employment growth rate data provided by 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

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

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

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

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 1.1 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.

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

Table 1.1: 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	0.1435	0.1359	-9.47
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	15.645	22.610	44.52

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.

The structural parameters that this calibration exercise is most sensitive to are the ones that govern the return to scale technology of the production function, δ and α. Table 1.2 shows the implied standard deviation, holding the persistence parameter fixed at $\rho^I = 0.95$. Stronger the decreasing returns to scale, larger is the implied volatility for the idiosyncratic TFP process.

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 informa-

Table 1.2: Implied standard deviation for the idiosyncratic TFP process - changing returns to scale parameters

	pre 1984	post 1984	% change
Average standard deviation (firm-level employment growth rate data)	0.4996	0.4730	-9.46
Returns to Scale $\delta + \alpha = 0.85$ $(\delta = 0.57, \alpha = 0.28)$			
Implied σ_u	0.2518	0.2384	-9.47
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	27.453	39.674	44.52
Returns to Scale $\delta + \alpha = 0.90$ $(\delta = 0.60, \alpha = 0.30)$			
Implied σ_u	0.1678	0.1589	-9.47
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	18.295	26.444	44.54
Returns to Scale $\delta + \alpha = 0.95$ $(\delta = 0.63, \alpha = 0.32)$			
Implied σ_u	0.0839	0.0795	-9.48
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	9.1474	13.23	44.63
Returns to Scale - Benchmark $\delta + \alpha = 0.896$ $(\delta = 0.64, \alpha = 0.256)$			
Implied σ_u	0.1746	0.1653	-9.46
Implied ratio $\frac{\sigma_u}{\sigma_\varepsilon}$	19.036	27.510	44.51

tion 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 1.3 summarizes the benchmark calibration.

1.5.2 Results

Figure 1.2 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 chapter 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

Table 1.3: Benchmark Parameters

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

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.

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 1.3. 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.

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

Figure 1.2: Impulse Response to an aggregate TFP shock

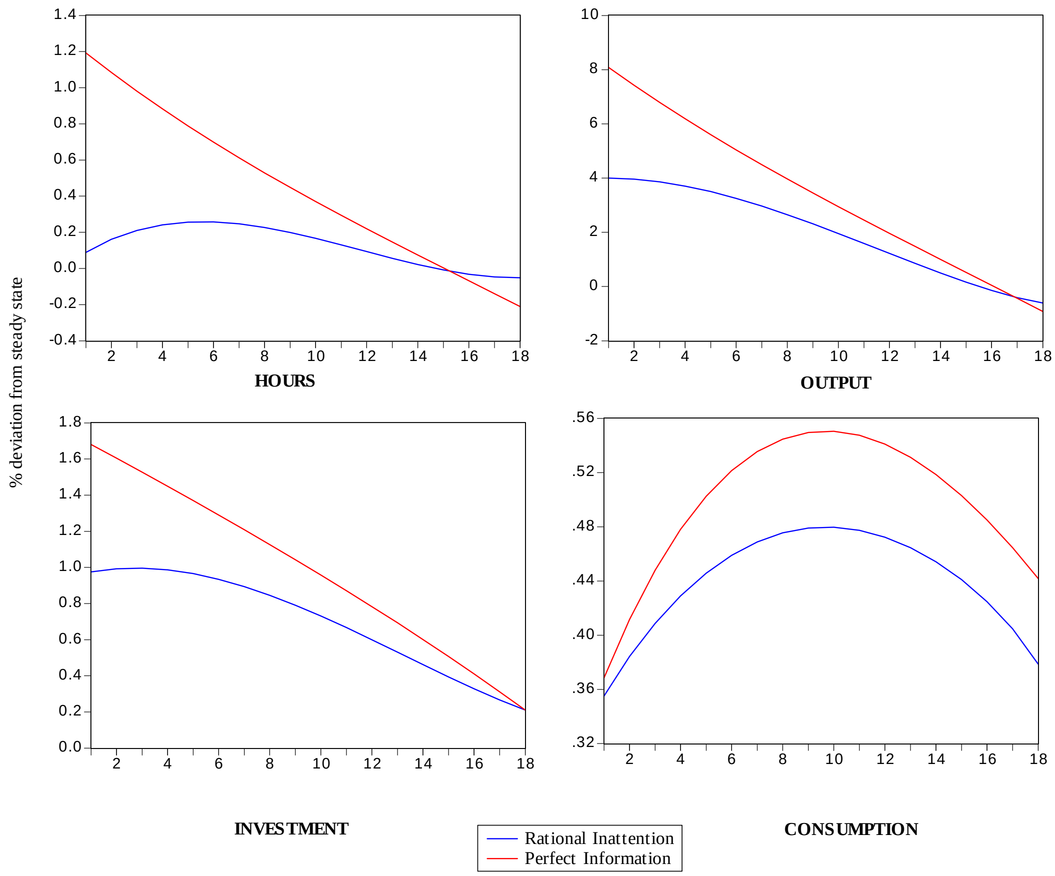


Figure 1.3: Impulse response of firm level input (labor and capital) choices to an innovation in idiosyncratic TFP

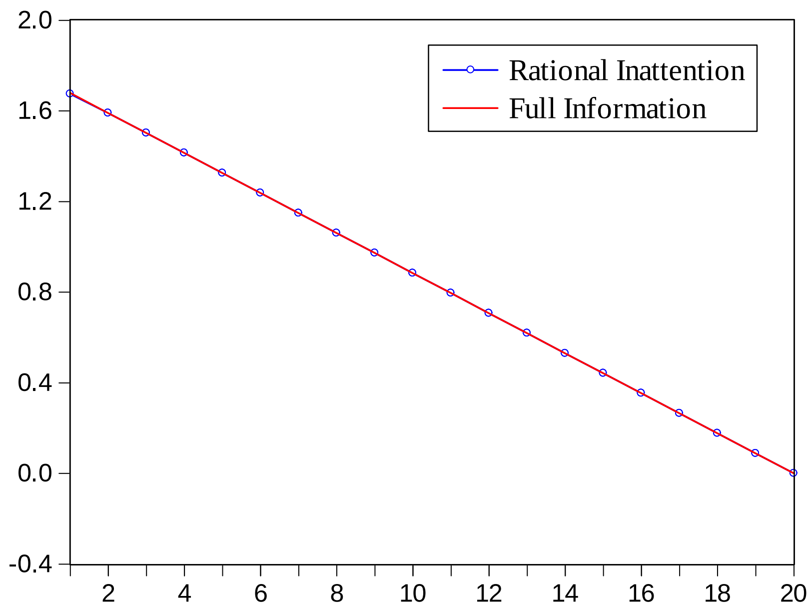


Figure 1.4: Business Cycle Statistics - Perfect Information vs Rational Inattention

		Cross Correlation of Output with : (Full Information)								
Variable	**SD (%)**	x(-4)	x(-3)	x(-2)	x(-1)	x	x(+1)	x(+2)	x(+3)	x(+4)
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

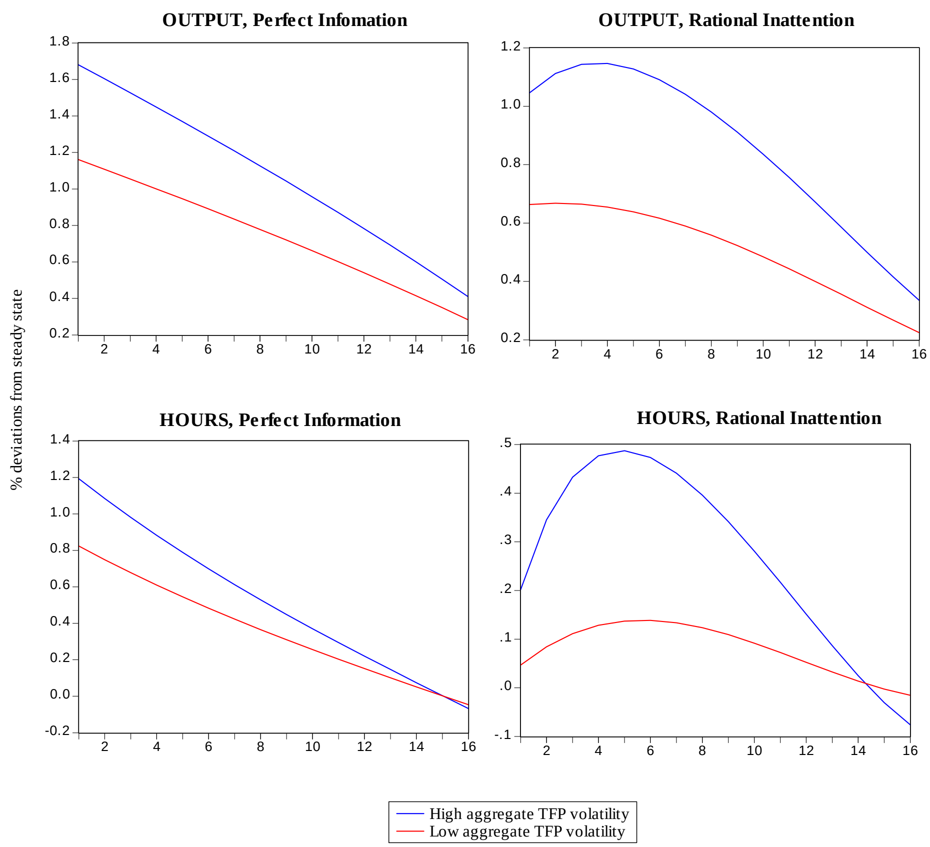
simulated data and compute the moments presented in Figure 1.4. Major differences between the perfect information and rational inattention models are observed in the volatility of aggregate variables. Note that given the simplifying assumption that the household sector in the economy has full information, 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.

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

Figure 1.5 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.

The magnitude of this amplification effect, which is the main result of this chapter, is summarized in Table 1.4. 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

Figure 1.5: Impulse Responses to an aggregate TFP shock across different TFP volatility regime and information structures



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 D for the proof.

Table 1.4: Great Moderation: Data versus RBC and Rational Inattention (RI)
(% 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**
$\sigma_\varepsilon(pre1984) = 0.0092, \sigma_\varepsilon(post1984) = 0.006, \frac{\sigma_\varepsilon(post1984)}{\sigma_\varepsilon(pre1984)} = 0.66$				

1.6 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 chapter. 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.

1.6.1 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.

$$\min E\left[\sum_{t=0}^{\infty}\beta^t(\frac{\hat{\pi}_{33}}{2}(\hat{k}_t^*-\hat{k}_t^F)^2+\frac{\hat{\pi}_{44}}{2}(\hat{l}_t^*-\hat{l}_t^F)^2+\hat{\pi}_{34}(\hat{k}_t^*-\hat{k}_t^F)(\hat{l}_t^*-\hat{l}_t^F))\right] \quad (1.52)$$

subject to

$$\hat{l}_t^F = \frac{1}{1-\alpha-\delta}(a_t-(1-\alpha)\hat{w}_t-\alpha\hat{r}_t) \quad (1.53)$$

$$\hat{k}_t^F = \frac{1}{1-\alpha-\delta}(a_t-\delta_t\hat{w}_t-(1-\delta)\hat{r}_t) \quad (1.54)$$

$$\hat{l}_t^* = E\left[\hat{l}_t^F|s_i^t\right] \quad (1.55)$$

$$\hat{k}_t^* = E\left[\hat{k}_t^F|s_i^t\right] \quad (1.56)$$

$$I(\{w_{t,}r_t,a\};\{s_{it}\}) \leq \kappa \quad (1.57)$$

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 1.5 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.

1.7 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

Table 1.5: Rational inattention (RI) without the attention allocation problem

(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**
$\sigma_\varepsilon(pre-1984)=0.0092, \sigma_\varepsilon(post-1984)=0.006, \frac{\sigma_\varepsilon(post-1984)}{\sigma_\varepsilon(pre-1984)}=0.66$				

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 1.6 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 second row in Figure 1.6.

Table 1.6 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

Figure 1.6: Impulse response of output and hours to an innovation in aggregate TFP across different idiosyncratic volatility regimes

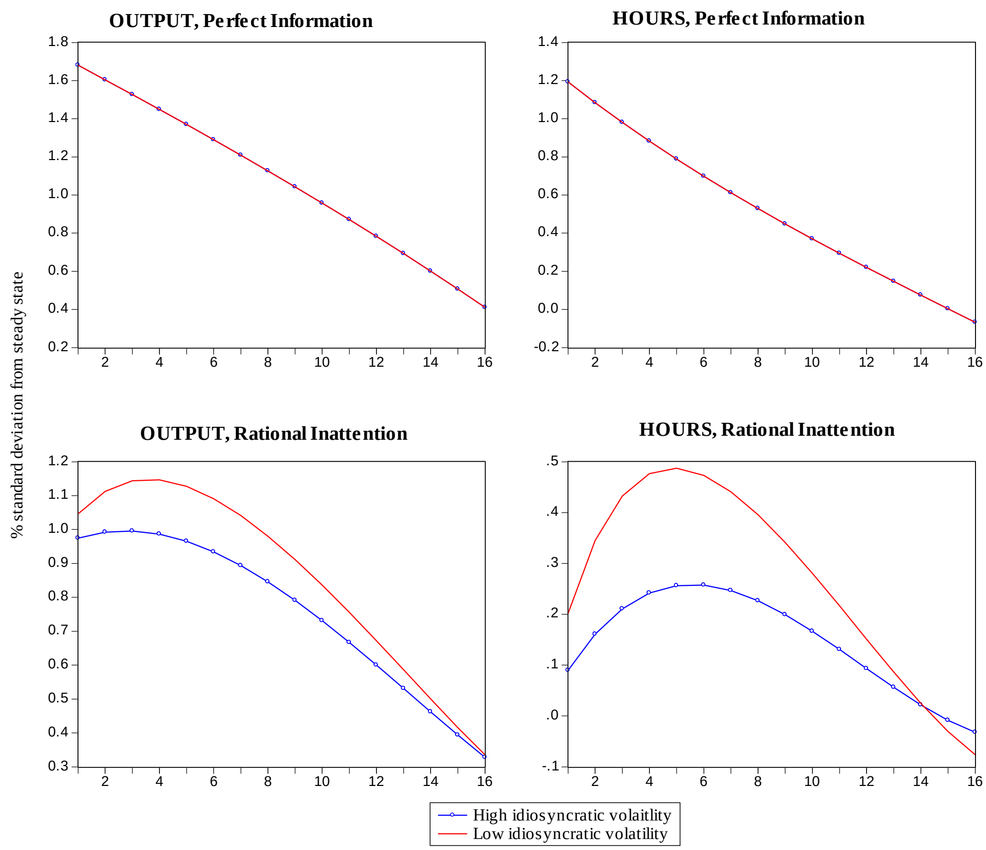


Table 1.6: 25% increase in idiosyncratic TFP volatility and no change in aggregate TFP volatility

(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
$\sigma_u(high) = 0.2242, \sigma_u(low) = 0.1746, \sigma_\varepsilon(high) = \sigma_\varepsilon(low)$				

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.

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. Abras (2010) documents the rise in earnings instability associated with a moderation in the aggregate as well as firm level activity.

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.

1.8 Sensitivity Analysis

In this section I examine the model's implication for different structural parameters such as the labor supply elasticity, the upper bound on information processing constraint, the assumed persistence parameter for the idiosyncratic TFP process, as well as different household preference specification.

1.8.1 Labor Supply Elasticity

Given the preferences used in the benchmark model, labor supply elasticity is defined as $\frac{1}{\psi}$. I compare the magnitude of the volatility amplification for different labor elasticity values. It must be noted that changes in ψ should be associated with changes in θ in order to maintain the same steady state value of time spent working that we observe in the data ($L^{ss} = 1/3$). All other parameters are kept unchanged.[15] As

[15]Not changing the other parameters does not have an effect on the steady state. The only steady state value that labor supply elasticity affects is hours (L).

the labor supply elasticity falls (as ψ increases) the volatility amplification effect does not change for hours but it falls for output and investment. More specifically, a 34% decline in aggregate TFP innovations leads to a decline in the volatility of aggregate output which varies in magnitude from 43% in the case of infinitely elastic labor supply to 38% in the case of unit elastic labor supply. The decline in aggregate investment volatility ranges from 46% to 40%. The decline in the volatility of hours remains roughly unchanged as labor supply elasticity changes.[16]

The intuition of why labor supply elasticity is important for the volatility amplification effect, can be found by looking at the equilibrium conditions in the labor market and how information is being processed. Equation (1.58) and (1.59) show the aggregate labor supply and labor demand equations, which I repeat here for convenience

$$\psi \hat{L}_t + \gamma \hat{C}_t = \hat{w}_t \tag{1.58}$$

$$\hat{L}_t^F = E\left[\tfrac{1}{1-\alpha-\delta}(a_t - (1-\alpha)\hat{w}_t - \alpha \hat{r}_t)|s^t\right] \tag{1.59}$$

Notice that given the assumption that households have full information, the labor supply decision does not depend on information processing constraints. Labor demand on the other hand, depends on the history and the set of signals that all firms in the economy receive. In equilibrium, labor demand and labor supply must equal each other, which implies that all the fluctuations in labor demand must be matched by

[16]Experiments show that volatility of hours is sensitive to changes in the labor supply elasticity alone. But in this exercise we must change the parameter that governs the disutility of labor as well in order to maintain the same steady state. As one can see from Table 1.7 lower labor supply need to be associated with higher labor supply disutility.

fluctuations in labor supply and vice versa. This adjustment is done via the wage rate in the economy, as can be seen from equation (1.58).[17] For a lower labor supply elasticity (high ψ), higher fluctuations in the wage rate would be required to reach the labor market equilibrium. Higher the volatility in wage rates, stronger the incentive of firms to pay more attention to the aggregate environment. And in fact, in all the numerical examples, lower the labor supply elasticity (higher ψ), higher is the amount of information processing capacity allocated to the aggregate environment. Hence, the reason why lower labor supply elasticity is associated with lower volatility amplification effect, which numerically is shown on Table 1.7, is that firms allocate more attention to the aggregate state. And firms allocate more attention to the aggregate state because wage rate volatility is bigger.

While changing the structural parameters the marginal value of information might change as well, questioning, in this case, how reasonable the assumed value of κ is. In all my experiments in this sensitivity analysis, the marginal value of information remains roughly the same. Hence, comparing two models with different structural parameters while maintaining κ unchanged is a valid exercise.

1.8.2 Upper Bound on Information Processing Capacity κ

As described in the benchmark calibration section of the model, the upper bound on the capacity to process information is chosen such that the loss in profits due to the

[17] The adjustment mechanism in the general equilibrium is more complicated than this, but focusing on the wage-channel captures the importance of labor supply elasticity.

Table 1.7: Robustness check - changing Labor Supply Elasticity (% standard deviation)

Series	**Output**	**Hours**	**Consumption**	**Investment**
$\psi = 0, \theta = 2.61, \kappa = 5$				
Pre-1984	2.01	1.22	0.62	9.29
Post-1984	1.14	0.53	0.41	5.06
Late/Early	**0.57**	**0.43**	**0.66**	**0.54**
$\psi = 0.1, \theta = 2.95, \kappa = 5$				
Pre-1984	1.89	1.01	0.59	8.66
Post-1984	1.07	0.41	0.39	4.67
Late/Early	**0.57**	**0.41**	**0.66**	**0.54**
$\psi = 0.3, \theta = 3.77, \kappa = 5$				
Pre-1984	1.70	0.70	0.56	7.70
Post-1984	1.02	0.31	0.38	4.41
Late/Early	**0.60**	**0.44**	**0.68**	**0.57**
$\psi = 1, \theta = 8.85, \kappa = 5$				
Pre-1984	1.50	0.35	0.52	6.65
Post-1984	0.93	0.15	0.36	3.98
Late/Early	**0.62**	**0.43**	**0.69**	**0.60**
$\sigma_\varepsilon(pre-1984) = 0.92, \sigma_\varepsilon(post-1984) = 0.6, \frac{\sigma_\varepsilon(post-1984)}{\sigma_\varepsilon(pre-1984)} = 0.66$				

lack of full information on the side of the firms is small enough not to induce them to invest in additional information processing capacity.[18] In Table 1.8 I report the model's results for different values of κ. As expected the higher the maximum capacity to process information, the lower the volatility amplification effect due to rational inattention. The reason is that higher κ leads to higher capacity being allocated to the aggregate TFP as well as idiosyncratic TFP. This means that firms will be able to observe the aggregate state more accurately and the aggregate dynamics approach those under full information.[19]

1.8.3 Persistence of the Idiosyncratic TFP Process

The assumed persistence parameter for the idiosyncratic TFP process is one of the important parameters that affect the allocation of attention by firms. As discussed by Maćkowiak and Wiederholt (2009a), changes in the persistence of an AR(1) process (keeping variance constant) have ambiguous effects on the amount of attention allocated to that variable. On one hand, a lower persistence, everything else equal, makes a process more difficult to track and hence it leads to more attention being allocated to it. On the other hand, a lower persistence may also increase or decrease the

[18] The idea is that what matters for firms profits is the idiosyncratic variables and not the aggregate ones. Hence this leads firms to allocate almost all of the information processing capacity to processing information about idiosyncratic variables. In this sense the mistakes firms make regarding the aggregate state have very little impact on their own profits *but* substantial impact on the aggregate dynamics.

[19] Please refer to Table 1.4 for comparison with the full information case. Within the same TFP-volatility regime, the higher κ is, the closer the volatility of aggregate variables is to those under full information.

Table 1.8: Robustness check - changing the upper bound of Information Processing Capacity

(% standard deviation)

Series	**Output**	**Hours**	**Consumption**	**Investment**
		$\kappa = 4.7$		
Pre-1984	1.75	0.80	0.60	7.84
Post-1984	0.95	0.22	0.40	3.92
Late/Early	**0.54**	**0.28**	**0.67**	**0.50**
		$\kappa = 4.9$		
Pre-1984	1.85	0.96	0.60	8.47
Post-1984	1.05	0.38	0.39	4.57
Late/Early	**0.57**	**0.40**	**0.65**	**0.54**
		$\kappa = 5.1$		
Pre-1984	1.93	1.06	0.60	8.90
Post-1984	1.11	0.48	0.39	4.95
Late/Early	**0.58**	**0.45**	**0.65**	**0.56**
		$\kappa = 5.3$		
Pre-1984	1.95	1.11	0.60	9.07
Post-1984	1.16	0.56	0.39	5.27
Late/Early	**0.59**	**0.50**	**0.65**	**0.58**
		$\kappa = 6.15$		
Pre-1984	2.15	1.38	0.60	10.21
Post-1984	1.34	0.80	0.39	6.26
Late/Early	**0.62**	**0.58**	**0.65**	**0.61**
$\sigma_\varepsilon(pre-1984) = 0.92, \sigma_\varepsilon(post-1984) = 0.6, \frac{\sigma_\varepsilon(post-1984)}{\sigma_\varepsilon(pre-1984)} = 0.66$				

marginal value of information, which leads to an increase or decrease in the attention allocation to that variable. In this model, lowering the persistence of the idiosyncratic shock while holding everything else constant leads to *less* attention allocated to the idiosyncratic shock.

In the calibration of the idiosyncratic TFP process, Table 1.1, different assumed persistence parameters lead to different implied volatilities as well. More specifically, lower persistence parameters are associated with lower implied volatilities.[20] Hence, in order to evaluate the effect of a lower persistence parameter, I have to use the implied volatility associating with it. Table 1.1 shows that the calibration exercise for $\rho^I = 0.5$ yields an implied standard deviation of 0.1537 and 0.1456 for pre-1984 and post-1984 periods respectively. Table 1.9 shows the results on aggregate volatility and the amplification effect of rational inattention for the benchmark calibration ($\rho^I = 0.95$) and for a less persistent idiosyncratic TFP process ($\rho^I = 0.5$). Results show that within a given volatility regime for the aggregate TFP (pre or post -1984), a lower persistence for the idiosyncratic TFP process (which is also associated with a lower volatility as well) leads to more attention being allocated to the aggregate state and less to the idiosyncratic one. This implies a better tracking of the aggregate environment and hence a greater volatility for each aggregate variable. Given the higher level of attention allocated to the aggregate conditions, the volatility amplification effect that rational inattention produces is lower in this case. A 34% reduction in the volatility of

[20] I note that because I am matching the volatility of the *growth rate* of employment as oppose to is *level*, the relationship between the persistence parameter and the volatility of the idiosyncratic TFP process is a positive one.

Table 1.9: Robustness check - Persistence of the idiosyncratic TFP process (% standard deviation)

Series	Output	Hours	Consumption	Investment
$\rho^I = 0.95, \sigma_u(pre-84) = 0.1746, \sigma_u(post-84) = 0.1653$				
Pre-1984	1.75	0.80	0.60	7.84
Post-1984	0.95	0.22	0.40	3.92
Late/Early	**0.54**	**0.28**	**0.67**	**0.50**
$\rho^I = 0.5, \sigma_u(pre-84) = 0.1537, \sigma_u(post-84) = 0.1456$				
Pre-1984	1.85	0.95	0.60	8.45
Post-1984	1.05	0.38	0.39	4.55
Late/Early	**0.57**	**0.40**	**0.65**	**0.54**

aggregate TFP innovations, leads to a 43% reduction in aggregate output volatility as opposed to the 46% reduction in the benchmark calibration. It is important to mention that direct estimates of plant-level and aggregate TFP shock persistence parameter (Cooper and Haltiwanger, 2006) point in the direction of higher idiosyncratic persistence as compared to the aggregate, which in this model works in favor of higher volatility amplification.

1.8.4 Different Household Preferences

Here I explore the implications that the form of household preferences has for the volatility amplification effect. I compare the results for the benchmark separable preferences versus the preferences assumed in Greenwood-Hercowitz-Hoffman (GHH, 1988).[21] 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:

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

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

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

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.

Table 1.10 reports the parameters used in the numerical solution for each preference specification. Labor supply elasticity for GHH preferences is $\frac{1}{\nu-1}$, $\nu > 1$ and for our benchmark preferences is $\frac{1}{\psi}$, $\psi > 0$. In this section's exercise I set ν such that it produces the same labor supply elasticity as in the benchmark model, that is, $\nu = \psi + 1$ while adjusting the parameters governing disutility of labor θ such that both models yield the same steady state results.

Table 1.11 shows the numerical results. First, within the same subperiod, GHH preferences lead to higher aggregate volatility. Second, as mentioned above rational inattention produces less of a volatility amplification effect in the case of GHH preferences. Third, there is a smaller asymmetry across aggregate variables in terms of the reduction in volatility as a response to a less volatile aggregate TFP process. This is a desirable feature since the benchmark model predicts strong counter factual results

Table 1.10: GHH and Benchmark Preferences - Parameters

Parameter	**Values**	**Description**
		Common Parameters
β	0.99	discount factor
γ	1	coefficient of relative risk aversion
d	0.02	depreciation rate
α	0.256	capital's share in output
δ	0.64	labor's share in output
κ	5	upper bound on information flow (bits)
ρ_A	0.95	persistence parameter for aggregate TFP process
ρ_I	0.95	persistence parameter for idiosyncratic TFP process
σ_ε (pre-1984)	0.0092	standard deviation of the innovation in aggregate TFP
σ_ε (post-1984)	0.006	standard deviation of the innovation in aggregate TFP
σ_u (pre-1984)	0.1746	standard deviation of the innovation in idiosyncratic TFP
σ_u (post-1984)	0.1653	standard deviation of the innovation in idiosyncratic TFP
		GHH Preferences
ν	1.3	labor supply elasticity: $\frac{1}{\nu-1}$
θ	1.76	the level of disutility of labor
		Benchmark Preferences
ψ	0.3	labor supply elasticity: $\frac{1}{\psi}$
θ	3.77	the level of disutility of labor

in terms of the disproportionate response of hours as compared to other aggregate variables. As shown in appendix D, the reaction of hours in terms of volatility reduction as compared to output will always be greater but the extent of this difference depends on the type of preferences being considered.

1.9 Endogenous Information Processing Capacity (κ)

In this subsection I explore the implications of rational inattention when the firms in addition to deciding how to allocate information, also decide how much information processing capacity they want to acquire. I assume that firms face a cost function $C(\kappa)$ when acquiring additional κ.

The attention allocation problem of the firms as described by equations (1.42) and (1.43), can be restated in terms of information flow allocated to aggregate versus idiosyncratic variables as opposed to signal-to-noise ratios. Let κ^A and κ^I, denote the amount of information allocated to the aggregate and idiosyncratic shock respectively. Any given pair $\{\kappa^A, \kappa^I\}$ is associated with the following signal-to-noise ratios: $\frac{\sigma_a^2}{\sigma_u^2} = 2^{\kappa^A} - 1$ and $\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2} = 2^{\kappa^I} - 1$. This comes from the information flow constraint, equation (1.43)

$$\underbrace{\frac{1}{2}\log_2\left(1+\frac{\sigma_a^2}{\sigma_u^2}\right)}_{\kappa^A} + \underbrace{\frac{1}{2}\log_2\left(1+\frac{\sigma_{ai}^2}{\sigma_\varepsilon^2}\right)}_{\kappa^I} \leq \kappa$$

It implies that choosing $\{\kappa^A, \kappa^I\}$ is the same as choosing the signal-to-noise ratios. The objective loss function that firms face due to imperfect information (1.42) can

Table 1.11: GHH vs Benchmark Preferences - Rational inattention (RI) versus standard RBC model

(percent standard deviations)

	GHH Preferences				*Benchmark Preferences*			
Series	**Output**	**Hours**	**Consumption**	**Investment**	**Output**	**Hours**	**Consumption**	**Investment**
RI (pre 1984)	2.24	1.57	1.42	6.70	1.70	0.70	0.56	7.70
RI (post 1984)	1.39	0.93	0.87	4.29	1.02	0.31	0.38	4.41
RI (late/early)	**0.62**	**0.59**	**0.61**	**0.64**	**0.60**	**0.44**	**0.68**	**0.57**
RBC (pre 1984)	2.55	1.96	1.67	6.90	2.11	1.27	0.56	10.14
RBC (post 1984)	1.66	1.28	1.09	4.50	1.38	0.83	0.36	6.63
RBC (late/early)	**0.65**	**0.65**	**0.65**	**0.65**	**0.65**	**0.65**	**0.65**	**0.65**

$\sigma_{\varepsilon}(pre-1984) = 0.92, \sigma_{\varepsilon}(post-1984) = 0.6, \frac{\sigma_{\varepsilon}(post-1984)}{\sigma_{\varepsilon}(pre-1984)} = 0.66$

be rewritten as

$$\min_{\kappa^A,\kappa^I} \frac{|\ \pi_{33}\ |}{2} \left\{ \left(\frac{1-\varphi}{1-\delta}\right)^2 \sigma_a^2 2^{-2\kappa^A} + \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{-2\kappa^I} \right\}$$

subject to

$$\kappa = \kappa^A + \kappa^I$$

Below I explore the implications of having a linear and a convex cost structure.

1.9.1 Linear Costs in Acquiring Information Processing Capacity

In addition to deciding how to allocate a given information processing capacity, firms also decide how much of this capacity to acquire. The firm's problem then becomes

$$\min_{\kappa^A,\kappa} \frac{|\ \pi_{33}\ |}{2} \left\{ \left(\frac{1-\varphi}{1-\delta}\right)^2 \sigma_a^2 2^{-2\kappa^A} + \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{-2(\kappa-\kappa^A)} \right\} - C(\kappa) \tag{1.60}$$

where I have substituted the constraint $\kappa^I = \kappa - \kappa^A$. Let's consider a linear cost structure $C(\kappa) = c\kappa$, where c is the marginal cost of acquiring an additional information processing capacity. The first order conditions for this problem are

$$2^{2\kappa^A} = (1-\varphi)\left(\frac{\sigma_a}{\sigma_{a_i}}\right) 2^{\kappa} \tag{1.61}$$

and

$$\ln(2)\ |\ \pi_{33}\ |\ \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{2(\kappa^A-\kappa)} = c \tag{1.62}$$

The first equation captures the attention allocation decision for a given capacity κ, and the second equation balances the marginal benefit and cost of acquiring additional information processing capacity. Looking for an interior solution, equations (1.61) and (1.62) will determine the optimal allocation of attention as well as the optimal amount of information processing capacity. The solution to this system of equations is

$$2^{\kappa} = \left[\frac{\ln(2) \mid \pi_{33} \mid}{c} \frac{(1-\varphi)}{(1-\delta)^2} \right] \sigma_a \sigma_{a_i} \tag{1.63}$$

and

$$2^{2\kappa^A} = \left(\frac{1-\varphi}{1-\delta} \right)^2 \left(\frac{\ln(2) \mid \pi_{33} \mid}{c} \right) \sigma_a^2 \tag{1.64}$$

There are two important outcomes when κ is endogenized assuming a linear cost structure. First, as equation (1.63) shows, the amount of optimal information processing capacity is an increasing function of the volatility of each shock.[22] This implies that, as each shock becomes more volatile (keeping the volatility of the other shock constant), it is optimal to increase the capacity to process information. Second, and most important, the optimal amount of information allocated to the aggregate shock is no longer a function of the ratio of idiosyncratic versus aggregate shock volatility. The optimal amount of attention now depends on own-shock volatility not on the relative volatility of shocks. This is important since it is in stark difference with the result obtained when κ was held fixed. The reason for such a result is the linear cost structure in obtaining new information processing capacity. In order to see why this is the case, let's focus on equations (1.61) and (1.63). When the volatility of the aggregate shock increases, there are two effects on the optimal level of attention

[22] At this point of the problem I haven't solved the fixed point problem yet (equilibrium φ), but as it will be shown later, $\varphi < 1$, which ensures a positive coefficient in equation (1.63).

allocated to the aggregate shock (κ^A) : first, firms would want to substitute capacity away from the idiosyncratic shock (equation (1.61)), and second, firms would also like to increase their total information processing capacity (equation (1.63)). I call the first the *capacity substitution effect* and the second, the *capacity acquisition* effect. In this case as I have shown before, an increase (decrease) in the volatility of the aggregate shock (keeping the idiosyncratic shock volatility constant), will lead to a higher (lower) level of information flow being allocated to the aggregate shock. The difference from the fixed-κ case is that all this increase comes from a higher overall capacity being acquired by the firm not due to a substitution of attention across states. Going back to equations (1.61) and (1.63), I explore the effect that a change in the volatility of the idiosyncratic shock has on the allocation of attention to the aggregate shock. By looking at equation (1.61) as the volatility of the idiosyncratic shock increases (σ_{a_i}), the attention allocated to the aggregate shock (κ^A) will tend to decrease since firms would want to substitute information from the aggregate to idiosyncratic shock. This captures the familiar substitution effect I explored the previous section, where κ was held fixed. In this case however, endogeneizing κ leads to an additional effect. When the volatility of the idiosyncratic shock increases, the overall capacity κ will also tend to increase, as can be seen from equation (1.63). In the case of a linear cost structure these two opposing effects cancel each other, leaving κ^A unchanged.[23] As σ_{a_i} increases, all the new acquired capacity is fully devoted to an increase in attention to

[23]Technically speaking, under a different cost structure, we would have a term $C'(\kappa)$ instead of c in equations (1.63) and (1.61).

idiosyncratic shocks and no change in the attention allocated to the aggregate shock. *Hence, under a linear cost structure, changes in the idiosyncratic environment have no effect on the macroeconomic environment.*

The equilibrium to this model is the solution to the fixed point problem between the initial guess (1.44) and the actual aggregate labor's law of motion.[24]

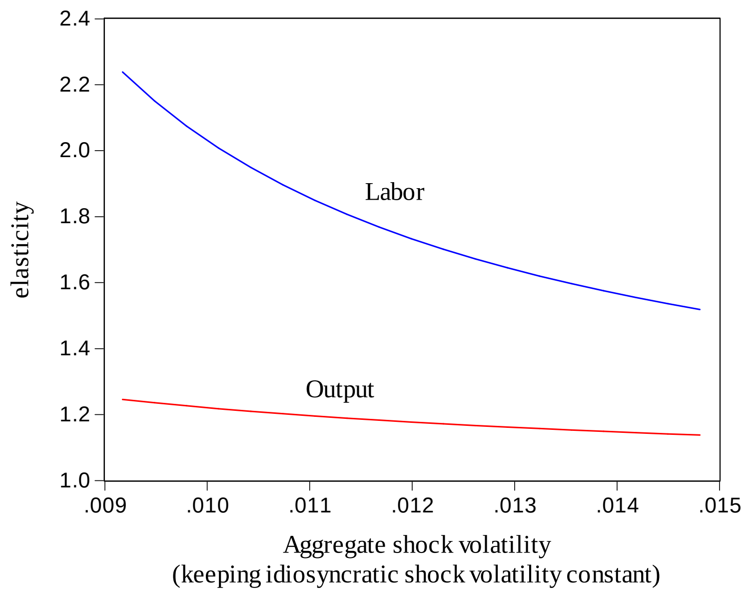
$$L = \frac{1-\varphi}{1-\delta}\left(1 - 2^{-2\kappa^A}\right)$$

The analytical solution to this problem, somewhat tedious, can be found in Appendix A. For expositional purposes I run a simple numerical example to show the effects of endogeneizing information processing capacity on the main focuses of this chapter, which is the elasticity of the aggregate shock volatility and aggregate outcome volatility.

Changes in the volatility of the aggregate shock are amplified even in the case of endogenous κ. That is, even though the model achieves a dichotomy between the idiosyncratic and the aggregate environment in the face of changes in the idiosyncratic volatility, it can still provide an amplification of the volatility of aggregate shocks. When the volatility of the aggregate TFP shock changes, the volatility of aggregate variables will still change by more. Figure 1.7 shows that for both aggregate variables, the elasticity of aggregate volatility with respect to aggregate shock volatility is greater than one, and that as the volatility of the aggregate shock gets larger, this

[24]This expression comes from aggregating individual labor-input responses of all firms under rational inattention.

Figure 1.7: Elasticity of aggregate volatility with respect to aggregate shock volatility. Linear cost in acquiring new information processing capacity.



elasticity falls. The decline in elasticity is due to the non-linear nature of the information flow constraint.[25] However, as previously shown, there will be no change in the volatility of aggregate variables if the volatility in the idiosyncratic shock changes.

1.9.2 Convex Costs in Acquiring Information Processing Capacity

In this section I experiment with a convex cost structure in the acquisition of information processing capacity and consider the implications that an endogenously determined κ has on the equilibrium. The problem the firm faces is represented by (1.60)

$$\min_{\kappa^A,\kappa} \frac{|\pi_{33}|}{2} \left\{ \left(\frac{1-\varphi}{1-\delta}\right)^2 \sigma_a^2 2^{-2\kappa^A} + \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{-2(\kappa-\kappa^A)} \right\} - C(\kappa)$$

where now $C'(.) > 0, C''(.) > 0$. For expositional purposes I choose $C(\kappa) = 2^\kappa$, simply because it allows for neater closed form expressions. The first order conditions for this version of the problem are[26]

$$2^{2\kappa^A} = (1-\varphi)\left(\frac{\sigma_a}{\sigma_{a_i}}\right)2^\kappa$$

and

$$|\pi_{33}| \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{2(\kappa^A-\kappa)} = 2^\kappa \tag{1.65}$$

Solving this system of equations leads to

$$2^\kappa = \left[|\pi_{33}| \frac{1-\varphi}{(1-\delta)^2}\right]^{\frac{1}{2}} \sigma_a^{\frac{1}{2}} \sigma_{a_i}^{\frac{1}{2}} \tag{1.66}$$

[25]See equation (1.43)

[26]The first order condition wrt to κ^A is the same and i repeat the equation here for completeness.

and

$$2^{2\kappa^A} = \left[\frac{(1-\varphi)^3}{(1-\delta)^2} \mid \pi_{33} \mid\right]^{\frac{1}{2}} \left(\frac{\sigma_a^3}{\sigma_{a_i}}\right)^{\frac{1}{2}} \tag{1.67}$$

Allowing for non-linear costs in acquiring information processing capacity, reconnects the macro and the microeconomic environment, similar to Section 1.4, where κ was held fixed. Changes in the idiosyncratic environment can affect the aggregate volatility. In order to see why this is the case let's focus on equations (1.61) and (1.66). The attention allocated to the aggregate shock (κ^A) is increasing in the volatility of this shock since both the *capacity substitution* effect and the *capacity acquisition* effect work in the same direction. As shown before such a result is true for a linear cost function as well. The difference between the two capacity acquisition technologies lies in the way that optimal allocation of attention to one shock reacts to changes in the volatility of the *other* shock. In this case, as the volatility of the idiosyncratic shock increases, there are two opposing effects on the optimal amount of attention allocated to the aggregate shock. As equation (1.61) shows, for a fixed κ there will be a tendency to decrease the attention allocated to the aggregate environment (κ^A). On the other hand, as equation (1.66) shows, an increase in the noise of any shock would make it optimal to increase the overall capacity (κ). Hence, on one hand, there is the *capacity substitution* effect that lowers the attention allocated to the aggregate shock and on the other hand, there is the *capacity acquisition* effect that increases this same attention. In the case of linear costs, these two effects cancel each other out. In the case of convex costs, on the other hand, the substitution effect is greater than the acquisition effect and hence as equation (1.67) shows, an increase in the volatility of the idiosyncratic shock lead to a reduction in the attention allocated to the aggregate shock. Hence, in this setting we are able once more to connect the idiosyncratic environment to the aggregate one.

Solving for the fixed point of this problem follows the same procedure as before. The two main results that were obtained in the fixed-κ case still hold in the case of convex costs. That is, there will be an amplification in the volatility of aggregate TFP shock as well as an impact of changes in the idiosyncratic noise onto the volatility of aggregate outcomes.

1.10 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. This chapter 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 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 chapter 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 1.7 of the chapter, 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.

Chapter 2

Welfare Cost of Anticipated Inflation in a Heterogeneous Agent Model

2.1 Introduction

This chapter examines the redistributive effects of monetary policy using a dynamic general equilibrium model with heterogenous agents. I study the long-run effects of inflation on output, consumption and welfare, as well as the distribution of wealth in the economy. Unlike in representative agent models, heterogeneity can potentially allow for beneficial effects of inflation. Increases in the growth rate of money supply can reduce wealth dispersion, increasing output and welfare.

This chapter builds on the two-sector search-theoretic model of Lagos and Wright (2005), which provides micro-foundations for money. One sector is characterized by decentralized trade where trading partners are matched randomly. This sector incorporates search and information frictions, which make money the essential medium of exchange. In addition to decentralized trade, agents have access to another sector where Walrasian markets operate and agents produce and trade goods, and adjust their money balances. Money is the only asset in the economy. It is a medium of exchange as well as the only store of value. For tractability purposes, preferences are assumed to be quasi-linear, eliminating any wealth effects in the demand for money and making the distribution of money holdings at the end of the centralized market session degenerate. This eliminates the extreme degree of market incompleteness. The sector where centralized trade occurs basically insures against all trading shocks that agents face in the decentralized market. All agents choose the same level of money holdings to carry into the next period. The Lagos and Wright (2005) model provides a tractable way of evaluating the welfare cost of inflation in an environment where the role of money is an endogenous outcome of search and information frictions. Due to its simplifying assumptions, which lead to a degenerate distribution of money holdings, it cannot be used to analyze the redistributive aspect of inflationary policies or their impact on the real economy and welfare. In order to study these aspects of inflation I augment the Lagos and Wright model in two ways. First, I introduce heterogeneity in discount factors and second, I allow the presence of productive capital in the economy.[1] I evaluate each element systematically, first by solving a model where money is the only asset in the economy but where agents differ in their discount factors, and second by solving a model where I allow for productive capital in the economy. This type of ex-ante heterogeneity in a Lagos and Wright framework

[1] One could introduce other forms of heterogeneity, such as, heterogeneity in preferences, productivity, etc.

provides a non-degenerate distribution of money while keeping the model tractable.[2] Agents have either high or low discount factors. The result is a two point distribution of money holdings and inflation tax has redistributive effects. The following results emerge from the first model. More patient agents hold more money than impatient ones. As long as money is being injected in the economy using lump-sum transfers, an increase in the growth rate of money supply has two effects: A direct effect of redistributing wealth from the rich to the poor, since the poor will have less than average money holdings, and an indirect effect of reducing real money balances for both agents. The sensitivity of the agent's money demand to inflation will be different for each type. In this model, the richer agents avoid the inflation tax faster than the poor agents. The net effect is that the direct redistributive effect of inflation in favor of the less wealthy is dominated by their weaker ability to evade the inflation tax. Next, I measure the welfare cost of inflation for each type of agent. The inflation tax seems to be affecting the less wealthy more than the wealthy agents, making inflation in this way a regressive tax. Erosa and Ventura (2002), using a different monetary model, reach the same conclusion.[3]

In the second model, I allow for the agents to accumulate human capital by allocating a fraction of their time to skill acquisition activities. Human capital can be used

[2]The quasilinearity assumption in the LW framework eliminates all the heterogeneity in money holdings that would emerge from trading shocks that agents face in the decentralized sector, but not any type of ex-ante heterogeneity such as heterogeneity in preferences, discount factors or other structural parameters.

[3]Erosa and Ventura (2002) build a model where agents hold money because buying goods with credit is costly. They show how inflation is a regressive consumption tax because wealthy agents have access to financial markets which allow them to avoid the inflation tax.

in productive activities in the decentralized and centralized sectors of the economy. This model allows me to examine the effect of inflation in a richer environment. The presence of another capital in an heterogenous agents environment can provide an additional channel for redistributive effects of changes in the growth rate of money supply. This channel is the economy-wide price of effective labor, in this case the wage rate. Firms in the centralized market produce an homogenous product using aggregate effective labor. The wage rate in the economy depends on the returns to scale technology. It can be constant, in the case of constant returns to scale or depend on average effective labor, in the case of a decreasing returns to scale production technology. In the latter case, the wage rate provides a redistributive channel for inflation other than the lump-sum transfer injections of money supply by the central bank. The results of this model show that for constant returns to scale technology, increases in the money growth rate lead to a reduction in aggregate production, consumption and human capital accumulation, as well as hours of work and time spent in skill acquisition. It also leads to an increase in the dispersion of wealth and skill-level (human capital). The welfare analysis shows that, as in the first model, the poorer agents (the less patient) suffer more from the inflation tax than the richer agents do. In the case of decreasing return to scale technology of production in the centralized market, the wage will depend on economy-wide average effective labor. In this case the model predicts a reversal of the previous results. Aggregate consumption, human capital and time spent in skill acquisition increase with a higher money growth rate, whereas aggregate hours of work decline. Dispersion in wealth and human capital falls. In this version of the model, patient agents hold less money and accumulate less human capital in the steady state as compared to the impatient ones. Welfare cost analysis shows that richer agents suffer most from inflation tax and poorer agents can actually benefit from inflation. In this scenario the Friedman rule is not the optimal policy.

Papers by Berentsen et al.(2005), Molico (2006), Bhattacharya et al. (2005), Berentsen and Strub (2009) have built on the Lagos and Wright (2005) framework to examine the redistributive effects of inflation. Bhattacharya et al. (2005) examine the redistributive effects of inflation in a very similar framework to the first model of this chapter, where money was the only asset in the economy. Berentsen and Strub (2009) in a similar model to Bhattacharya et al. (2005) study alternative institutional arrangements for the determination of monetary policy in a search-theoretic setting where agents differ in terms of preferences. The main difference between the first model in this chapter and Bhattacharya et al. (2005), is that the later assumes a different type of heterogeneity. Agents in their model are different in their consumption preferences as opposed to having different discount factors. The resulting effect of inflation on welfare is also different. Bhattacharya et al. (2005)'s results show that relatively richer agents suffer more from inflation as opposed to the less wealthy. I obtain the opposite result. In this model the ability of the wealthier agents to evade the inflation tax dominates the transfer from the rich to the poor that lump-sum injections of money provide. Molico (2006) studies the effects of money growth in a heterogenous agents model. The author also uses a search-theoretic model of money, but not of the Lagos and Wright (2005) type. Molico (2006) results show that for low inflation rates, lump-sum transfers of money compress the distribution of wealth and improve welfare. The opposite is true for higher inflation rates. Their heterogeneity is an endogenous one. Unlike in my first model, agents in Molico (2006) hold different amounts of money because of the history of trading shocks they face.[4]

Typically, search-theoretic models of money consider environments where money is the only asset in the economy. It is a medium of exchange as well as the only store of value. The first attempt to introduce another asset in a Lagos and Wright framework, namely physical capital in the centralized market (CM), was by Aruoba and Wright (2003). Physical capital was not introduced into decentralized market (DM) production since claims on physical capital would compete with money as a medium of exchange and potentially dominate it in a rate of return sense.[5] With physical capital being used for production only in the CM, the model dichotomizes. That is, one can solve for the DM production path and CM production and capital accumulation separately, with inflation having no effect on the latter. Money is super-neutral in terms of capital accumulation and CM production. Aruoba, Waller and Wright (2008) extend the Aruoba and Wright (2003) model by allowing physical

[4]Molico (2006) does not make the quasilinearity assumption that Lagos and Wright as well as this model does. Hence, agents depending on their trading shock history will end up with different amounts of money in each period.

[5]Lagos and Rochetau (2008) address the co-existence of money and capital as media-of-exchange.

capital to be used for productive purposes in the DM. This breaks the aforementioned dichotomy, so changes in the money growth rate have an effect in production and capital accumulation in the CM as well. The second model in this chapter differs from Aruoba, Waller and Wright (2008), in that it augments the search-theoretic model of Lagos and Wright (2005) with the decision to accumulate another type of capital, human capital, which is used for productive purposes in both markets. I also allow for heterogeneity in discount factors. In this setting, I can explore the effects of inflation on aggregate output, human capital accumulation and welfare as well as the distribution of wealth.

To my knowledge, Molico and Zhang (2005) is the only paper using search models of money that allows for a portfolio allocation decision in a heterogenous agents model by allowing agent to accumulate both money and capital (storable goods in their model).[6] Their results show that a moderate rate of monetary expansion can lead to an increase in steady-state aggregate output, aggregate consumption, capital accumulation, and welfare. Also, the average fraction of time spent working might decrease. My model is different from Molico and Zhang (2005) in that I consider an ex-ante heterogeneity in discount factors and human capital as opposed to tangible forms of capital. Their results are similar to the ones I obtain in the second model in the case of decreasing returns to scale technology of production.

[6]Even though their model has a two-sector economy, the sectors play different roles from the sectors in the Lagos and Wright (2005) model.

This chapter is organized as follows. In section 2.2 I solve a Lagos and Wright (LW) model which allows for heterogeneity in discount factors. In Section 2.4 I augment the previous model to allow for human capital accumulation.

2.2 A Search Economy: Money is the Only Asset in the Economy

There is a [0,1] continuum of infinitely lived agents operating in a Lagos and Wright type of economy. Time is discrete. Each period consists of two subperiods. It is also assumed that there are two types of goods, a special good produced and traded in the first subperiod and a general good traded in the second. In the first subperiod, which I will call the Decentralized Market (DM), agents trade in pairwise meetings. Agents receive a trading shock at the entrance of the DM. An agent can be in one of the following states: she can consume but not produce or produce but not consume. I assume that there is no double coincidence of wants (without loss of generality), so that agents refrain from bartering and I can focus on single coincidence meetings. Such search frictions generate endogenously the existence of an additional object called *money*, which enables trade in the DM and which cannot be consumed or produced by any agent. Money here is an intrinsically useless, non-perishable object used as a medium of exchange. In the second subperiod, the Centralized (Walrasian) Market, agents produce and consume a general good. They can transform one unit of labor into one unit of the general good. Similar to the Lagos and Wright model, I assume that preferences in the CM are quasilinear. Such an assumption implies that independently of the trading shock during the DM, agents exit the centralized market (CM) with the same level of money holdings. Achieving a degenerate distribution of money holdings at the end of each period increases the tractability of the model but it comes at the cost of ignoring the differential impact that inflation has in the economy. In this model, I allow agents to be different in their discount factors, which allows me to examine the redistributive effects of inflation while keeping the model tractable. I assume there is a monetary authority, namely the central bank, which injects money in the CM via lump-sum transfers denoted by τ. Money supply evolves via $M_t = (1+z)M_{t-1}$and $\tau = zM_{t-1}$. I start by examining the problem of an agent in the CM. The agent chooses consumption (X) of the general good, hours of work (H) and next period's amount of money holdings (m'), which maximizes

$$W_\alpha(m) = \max_{X,H,m'} \{U(X) - H + \beta(\alpha)V_\alpha(m')\} \tag{2.1}$$

subject to

$$X = \phi(m - m' + \tau) + wH \tag{2.2}$$

where ϕ is the units of consumption good per unit of money (inverse of price level), w is the wage rate[7], $W_\alpha(m)$ is the value function of a type α entering the CM with money m and $V_\alpha(m)$ is the value function of type α agent entering the DM with

[7]We can think of an environment where there are firms in the CM that employ only labor using linear production technology, which implies a constant wage rate. We set $w = 1$ for now.

money m. After substituting the budget constraint into the value function I get

$$W_\alpha(m) = \max_{X,m'}\{U(X) - [X - \phi(m - m' + \tau)] + \beta(\alpha)V_\alpha(m')\} \tag{2.3}$$

The first order conditions for the CM problem are

$$U'(X) = 1 \tag{2.4}$$

$$\phi = \beta(\alpha)V'_\alpha(m') \tag{2.5}$$

The envelope condition is

$$W'_\alpha(m) = \phi \tag{2.6}$$

Every agent consumes the same X, and the decision for the next period's money holdings is independent of this period's money holdings, but it does depend on the agents type α. Hence, agents of the same type will exit the CM with the same money holdings. $F(m_\alpha) = G(\alpha)$. This means that the distribution of money holdings is degenerate conditional on types.

In the DM agents come together in pairwise meetings. An agent can be in one of three possible situations. She can receive a consumption shock with probability σ and hence be a buyer, she can receive a production shock with the same probability σ and be a seller, or with probability $1 - 2\sigma$ she can be neither a consumer nor a producer. The only possible trades are goods for money, since I assumed no double coincidence of wants (barter). Letting $V_\alpha(m)$ denote the value function of a type α individual entering the DM with m, I have

$$V_\alpha(m) = \sigma\int[-c(q(\tilde{m}))+W_\alpha(m+d(\tilde{m}))]dF(\tilde{m})+\sigma[u(q(m))+W_\alpha(m-d(m))+(1-2\sigma)W_\alpha(m)$$

where $q(\tilde{m})$ is the quantity produced by a seller, which depends only on money balances of the buyer, and $d(\tilde{m})$ is the payment received by the seller. I am assuming, as will be verified below, that the quantity produced in the DM (q) and the amount that will have to be paid in exchange for the product (d) depend only on buyer's and not on seller's money balances. The first term captures the value of being a seller in the DM. Since the quantity produced by a seller ($q(\tilde{m}\ )$) will depend on the buyer's money balances, I integrate over the type distribution of buyers in the economy. The second and last term capture the value of being a buyer and not trading in the DM respectively. The marginal value of carrying money balances in the DM is

$$V'_\alpha(m) = \sigma \int W'_\alpha(m + d(\tilde{m}))dF(\tilde{m}) + \sigma[u'(q(m)) + W'_\alpha(m - d(m)) + (1 - 2\sigma)W'_\alpha(m) \tag{2.7}$$

Using the fact that $W'_\alpha(m) = \phi$, equation (2.7) becomes

$$V'_\alpha(m) = (1 - \sigma)\phi + \sigma[u'(q(m))q'(m) + \phi(1 - d'(m)) \tag{2.8}$$

Following LW, I assume that terms of trade are determined by generalized Nash Bargaining, where θ is the buyer's bargaining power. This problem is as follows

$$\max_{q,d\leq m}[u(q) + W_\alpha(m - d) - W_\alpha(m)]^\theta[-c(q) + W_\alpha(\tilde{m} + d) - W_\alpha(\tilde{m})]^{1-\theta}$$

Surplus from trading for the buyer is $[u(q) + W_\alpha(m - d) - W_\alpha(m)]$ and surplus from trading for the seller is $[-c(q) + W_\alpha(\tilde{m} + d) - W_\alpha(\tilde{m})]$. Making use of the fact that for each type α, $W_\alpha(m + d) - W_\alpha(m) = d\phi$ I can rewrite the Nash Bargaining problem as follows:

$$\max_{q,d\leq m}[u(q)+d\phi]^{\theta}[-c(q)+d\phi)]^{1-\theta}$$

As it is the case in LW, in equilibrium $d = m$ must hold. The quantity of goods being produced and consumed in the DM, $q = q(m)$ is the solution of

$$m\phi = \frac{\theta u'(q)c(q)+(1-\theta)c_q(q)u(q)}{\theta u'(q)+(1-\theta)c_q(q)} = g(q) \tag{2.9}$$

$$g_q > 0, q = q(m), \quad q'(m) = \frac{\phi}{g_q}, \quad d'(m) = 1 \tag{2.10}$$

After substituting equation (2.9) into equation (2.8), I get:

$$V_{\alpha}'(m) = \phi[1-\sigma+\sigma\frac{u'(q)}{g_q(q)}] \tag{2.11}$$

Substituting equation (2.11), into equation (2.5) I can derive the equilibrium condition for this economy:

$$\phi_t = \beta(\alpha)\phi_{t+1}[1-\sigma+\sigma\frac{u'(q_{t+1})}{g_q(q_{t+1})}] \tag{2.12}$$

At steady state, real money balances, $m\phi$, are constant. The law of motion for money balances, $m_{t+1} = (1+z)m_t$, implies that the law of motion for prices follows $\phi_t = (1+z)\phi_{t+1}$. As a result, the steady state equation for this economy becomes:

$$1+z = \beta(\alpha)[1-\sigma+\sigma\frac{u'(q_{\alpha})}{g_q(q_{\alpha})}] \tag{2.13}$$

Proposition 2 *At the steady state, relatively more patient agents choose to hold more money and consume more of the DM goods.*

Proof. Implicitly from the steady state equation we have $q = q(\alpha)$ and one can show that $\partial q/\partial\alpha > 0$.

$$\partial q/\partial\alpha = -\frac{1+z}{\beta(\alpha)^2\sigma[\frac{u_{qq}g_q - u_q g_{qq}}{g_q^2}]} > 0$$

Given that $u_{qq} < 0$, $g_q > 0$, $g_{qq} > 0$ we have $u_{qq}g_q - u_q g_{qq} < 0$. From the first order condition of the Nash Bargaining problem equation (2.9) we have:

$$\partial m/\partial\alpha = -\frac{-g_q\frac{\partial q}{\partial\alpha}}{\phi} > 0$$

■

There are two types of heterogeneity at the beginning of the CM, the trading shock in the DM and our imposed heterogeneity in discount factors. The money holdings of the agent entering the CM will depend on whether she was a buyer, a seller or no trade occurred. Looking at the budget constraint (2.2), there will be a variation in the hours of work as well. Using the fact that in steady state $\tau = zM$ and $m'_\alpha = (1+z)m_\alpha$, each type can find oneself in any of the three situations:

$$H_\alpha = \begin{cases} X - \phi(0 - (1+z)m_\alpha + zM) & \text{if previously a buyer in the DM, w/p } \sigma \\ X - \phi(m_\alpha + m_{\tilde{\alpha}} - (1+z)m_\alpha + zM) & \text{if previously a seller in the DM, w/p } \sigma G(\alpha) \\ X - \phi(m_\alpha - (1+z)m_\alpha + zM) & \text{if previously no trade, w/p } 1 - 2\sigma \end{cases} \tag{2.14}$$

where $M = \int m_\alpha dG(\alpha)$ is the average money balances in the whole economy. Given the quasilinear preferences, the level of next period's money balances is the same for agents with the same discount factor. Hence, CM labor effort absorbs the trading shock in the DM. In order to carry out the welfare analysis, I denote welfare for a type α agent as the sum of the expected steady state utility in the DM and in the CM

$$(1-\beta(\alpha))V_{\alpha} = \sigma[u(q_{\alpha}) - c(q_{\alpha})] + U(X) - \bar{H}_{\alpha} \tag{2.15}$$

where $\bar{H}_{\alpha}$ is the expected hours of work for type α computed using (2.14).

$$\bar{H}_{\alpha} = X - \phi(\sigma + z)(M - m_{\alpha}) \tag{2.16}$$

The welfare function then becomes

$$(1-\beta(\alpha))V_{\alpha} = \sigma[u(q_{\alpha}) - c(q_{\alpha})] + U(X) - X + \phi(\sigma + z)(M - m_{\alpha}) \tag{2.17}$$

I am interested in how higher rates of money growth rate affect welfare, hence

$$\frac{\partial(1-\beta(\alpha))V_{\alpha}}{\partial z} = \sigma\left[u'(q_{\alpha})\frac{\partial q_{\alpha}}{\partial z} - c'(q_{\alpha})\frac{\partial q_{\alpha}}{\partial z}\right] - \frac{\partial \bar{H}_{\alpha}}{\partial z} \tag{2.18}$$

The first term denotes the effect of a higher money growth rate on expected utility in the DM, whereas the second term denotes the impact on expected CM utility. Since the DM expected utility depends only on individual variables, it can only capture the rate of return effect of higher inflation. The second term reflects how expected utility in the CM is affected by a higher money growth rate. It is this term that captures interesting redistributive effects as we will show below. The change in the expected hours of work for each agent as a response to the higher money growth rate is:

$$\frac{\partial \bar{H}_{\alpha}}{\partial z} = -\left[\phi(M - m_{\alpha}) + (\sigma + z)\left(\frac{\partial \phi M}{\partial z} - \frac{\partial \phi m_{\alpha}}{\partial z}\right)\right] \tag{2.19}$$

The right hand side of equation (2.19) captures the two different redistributive aspects of inflation. The first term denotes the static redistributive effect. That is, assuming prices are fixed (agents have not yet adjusted to inflation), one's position in the distribution of money holdings determines whether one benefits or not from a higher money growth rate. Agents holding less than average money holdings will have to work less in the CM. The second term, denotes the dynamic redistributive effect. This term captures the differences in the responsiveness of money demand to inflation. That is, the degree to which different agents evade the inflation tax. If agents holding less than average money balances, have a stronger response to changes in the inflation rate, then both these redistributive effects work in the same direction, and an increase in the money growth rate would lead to a redistribution of wealth from the rich to the poor. If the opposite is true, that is, if agents holding more than average money balances, have a stronger response to inflation, then, the net result of redistribution will depend on which term dominates the other. Given that $\frac{\partial \phi M}{\partial z} < 0$, $\frac{\partial \phi m_\alpha}{\partial z} < 0$, I can express this term as : $(\sigma + z)(|\frac{\partial \phi m_\alpha}{\partial z}| - |\frac{\partial \phi M}{\partial z}|)$ and equation (2.19) becomes

$$\frac{\partial \bar{H}_\alpha}{\partial z} = -\left[\phi(M - m_\alpha) + (\sigma + z)\left(|\frac{\partial \phi m_\alpha}{\partial z}| - |\frac{\partial \phi M}{\partial z}|\right)\right] \tag{2.20}$$

Below, I provide an analytical example.

Example *Suppose* $u(q) = \ln q,\ c(q) = q,\ \theta = 1 \implies g(q) = c(q) = q$. *Then, the equilibrium quantities of the DM good and money holdings for each type are:*

$$q = \frac{\sigma\beta(\alpha)}{1 + z - \beta(\alpha)(1-\sigma)} \tag{2.21}$$

$$m = \frac{1}{\phi}[\frac{\sigma\beta(\alpha)}{1 + z - \beta(\alpha)(1-\sigma)}] \tag{2.22}$$

where, $q(\alpha_H) > q(\alpha_L), m(\alpha_H) > m(\alpha_L)$.*Consider an economy consisting of only two types of agents, the patient* (β^H) *and the impatient* (β^L).*Both types have equal mass. The aggregate amount of money in the economy is then defined as* $M = \frac{1}{2}m_L + \frac{1}{2}m_H$. *After inserting equations (2.21), (2.22) into equation (2.16) I get:*

$$\bar{H}_L = X^* - \frac{1}{2}(\sigma + z)\left(\frac{\beta_H\sigma}{1 + z - \beta_H(1-\sigma)} - \frac{\beta_L\sigma}{1 + z - \beta_L(1-\sigma)}\right)$$

One can show that :

$$\frac{\partial \bar{H}_L}{\partial z} > 0, \frac{\partial \bar{H}_H}{\partial z} < 0$$

In this example, changes in the money growth rate generate a transfer from the poor to the rich. Bhattacharya et al. (2005) examine the redistributive effects of inflation in a similar environment, where the agents are assumed to be heterogeneous in their consumption preferences. The authors obtain the opposite results in terms of the direction of the transfer. It is important to note at this point that the type of heterogeneity affects the direction of redistribution generated by inflation in the economy. In the next section I provide a numerical solution to a version of the model that does not allow for an analytical solution. The same results hold. That is, in a search economy, where agents are heterogenous in their discount factors, changes in the money growth rate generate a transfer from the poor to the rich.

The intuition for the above result can be found in the following tax/transfer argument. The real transfer each agent receives is

$$TR = \phi z M \implies \partial TR/\partial z = \phi M + \frac{\partial \phi M}{\partial z}$$

The real inflation tax each agent incurs is

$$IT = \phi z m_\alpha \implies \partial IT/\partial z = \phi m_\alpha + \frac{\partial \phi m_\alpha}{\partial z}$$

The difference between the two expressions is the net effect of a higher money growth rate z. The first term in each expression shows that before agents adjust to the new higher prices they benefit from a higher rate of monetary expansion as long as their money holdings are below average. The second term in each of the above expressions shows how fast the real transfer is falling and by how fast agents are able to evade inflation tax. The scenario in this chapter is such that for the poor agents the rate at which the real transfer is falling is bigger than the rate at which they are evading the inflation tax.

It is important to note that the aggregate hours of work in the CM, H, remain constant as the money growth rate changes. That is, the redistributive effects of inflation that allow patient agents to work less in the CM are exactly offset by the increase in working hours for impatient agents. In this respect, monetary policy is neutral in the CM. Nevertheless, redistributive effects have a significantly different impact on the welfare cost of inflation for each group of agents, as I will show in the next section.

2.3 Numerical Solution of the First Model

Here I provide some simulation results for parameter values that do not allow for analytical solutions that are easy to read. My parametrization (Table 2.1)and functional form choice follows that of Lagos and Wright (2005).

$$U(x) = B\log(x)$$

$$DM : u(q) = \frac{(q+b)^{1-\eta} - b^{1-\eta}}{1-\eta}$$

$$c(q) = q$$

In Figure 2.1, I decompose the impact of inflation on welfare among different subperiods. In the last row, we see how inflation affects welfare at the DM at the steady state. Expected utility is falling for both types in the DM. This basically reflects the inflation tax argument of monetary expansion on DM activity. The erosion of the purchasing power of money leads to lower money demand and hence less trade in the DM. Expected utility in the CM for different types of agents on the other hand, moves in opposite directions.

I compute the welfare cost of moving from 0% to 10% inflation in Table 2.2 for different parametrizations of the model. The low type agents, the impatient and the

Figure 2.1: Welfare implications of expansionary monetary policy

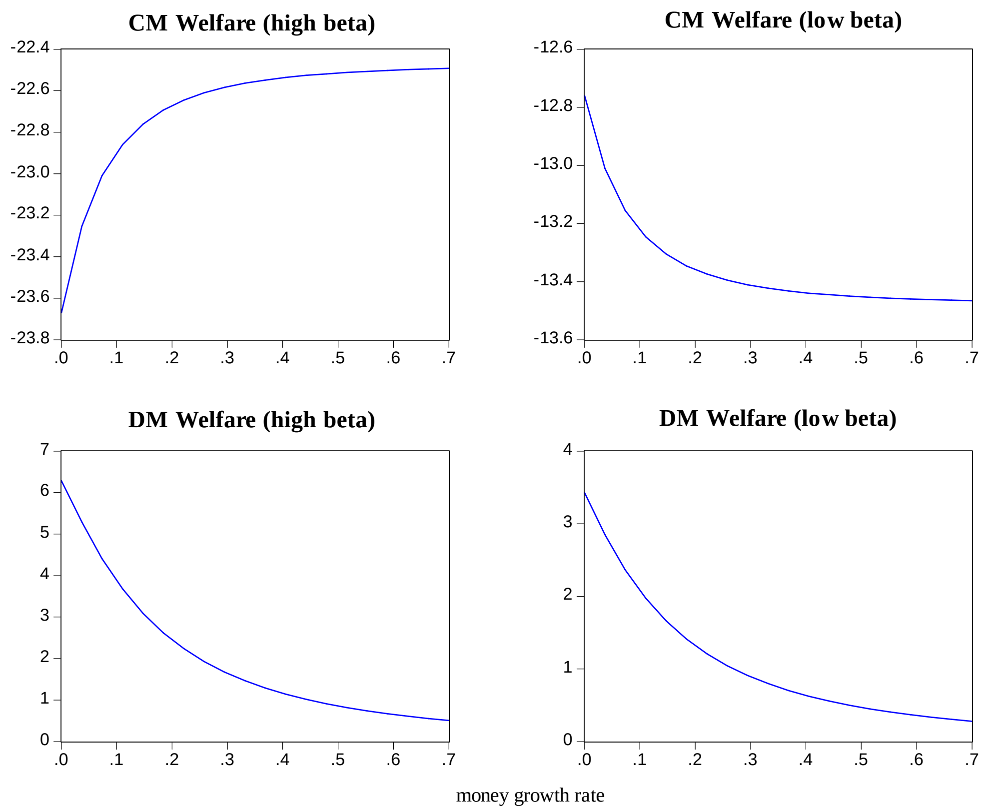


Table 2.1: Benchmark Parameter Values

Parameter	**Value**	**Description**
η	0.30	coefficient of risk aversion
θ	0.50	buyer's bargaining power
σ	0.50	probability of a bilateral meeting
B	1.91	constant
b	0.001	constant
ω_L	0.5	share of β_L types
β_L	0.94	discount factor for low-types
β_H	0.9615	discount factor for high-types

poor, suffer more from expansionary monetary policies relative to high types. The difference between the welfare cost of inflation between the two types increases as the degree of heterogeneity, the distance between the discount factors, increases. In an economy populated by agents with the same (high type) discount factor, the welfare cost of inflation will always be positive. If an economy is populated by both types of agents, then the welfare cost of inflation for the high type decreases and can also become negative. Redistributive effects of inflation from low to high types can be strong enough for inflation to be beneficial for a group of individuals. I obtain welfare benefits for high types when the difference in discount factors or the share of impatient agents is big enough (last three columns in Table 2.2).

Table 2.2: Welfare cost of moving from 0% to 10% inflation

(% consumption)

Cases	**Case 1**	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
Low Type	4.35	1.95	5.27	3.23	5.20	2.87
High Type	2.43	0.91	0.32	-1.56	-0.11	-2.23
Degenerate Distribution($\beta_H = \beta_L$)	3.52	1.44	3.52	3.52	3.52	3.52

Case 1 refers to the LW calibration and this model's benchmark parametrization, Table 2.1.

Case 2 refers to the LW calibration with $\theta = 1$. Case 3 refers to Case 1 with $\beta_L = 0.85$

Case 4 refers to Case 3 with $\omega_L = 0.8$ (higher share of the impatients)

Case 5 refers to Case 1 but with $\beta_L = 0.8$. Case 6 refers to Case 5 with $\omega_L = 0.8$

2.4 An Augmented Search economy: Money and Human Capital

In this section, I augment the Lagos and Wright model in yet another dimension. Agents in the economy can decide to accumulate another type of capital, human capital, which is used for productive purposes in both markets. Agents continue to be heterogeneous in discount factors. In this setting, I can explore the effects of inflation on aggregate output, human capital accumulation and welfare as well as the distribution of wealth

2.4.1 Model

Let $W_\alpha(m, h)$ and $V_\alpha(m, h)$ be the value functions for an agent in the CM and DM respectively, holding m dollars and h units of human capital. Human capital in this model is accumulated via time invested in education or in any knowledge acquiring activity. During the CM, agents rent their effective labor to firms which employ only labor. They choose what fraction of time to spend on accumulating human capital u or on market activity n as well as choosing future money balances and how much of the general good to produce. The CM problem for an agent becomes

$$W_\alpha(m, h) = \max_{x,u,n,m',h'} \{U(x) + A(1 - u - n) + \beta(\alpha)V_\alpha(m', h')\} \tag{2.23}$$

subject to

$$x = \phi(m - m' + \tau) + whn + \pi \tag{2.24}$$

$$h' = (1 - \delta^h)h + f(u, h) \tag{2.25}$$

$$\lambda_t(h' - (1 - \delta^h)h) \geq 0 \tag{2.26}$$

where w is the wage per unit of effective labor (hn), n is the fraction of time spent in market activities, u is the fraction of time invested in education, $1 - u - n$ is leisure. $f(u, h)$ captures the production function for human capital, which depends on time invested in education and on the level of current human capital. Firms operate under a one-input (effective labor) technology

$$Y = \left[\int n_\alpha h_\alpha dG(\alpha)\right]^\gamma$$

The parameter γ captures the returns to scale of the production technology. In the case of decreasing returns, profits generated by firms are distributed to consumers as π. A constant returns to scale production technology in the CM, implies a constant wage rate $(w = 1)$, whereas a decreasing returns to scale technology implies that the wage rate will depend on the economy-wide average effective labor. In this chapter, I will experiment with decreasing returns and constant returns to scale technology. This will turn out to be an important distinction in terms of the qualitative results of this chapter, since decreasing returns to scale technology introduces general equilibrium effects. After substituting n from the budget constraint (2.24) and u from (2.25) into

(2.23) I get[8]

$$W_\alpha(m,h) = \max_{x,m',u} \{U(x)+A-Au-A[\frac{x-\phi(m-m')-\pi}{wh}]+\beta(\alpha)V_\alpha(m',(1-\delta^h)h+f(u,h)\} \tag{2.27}$$

First order conditions for the CM problem are

$$x: \qquad U(x) = \frac{A}{wh} \tag{2.28}$$

$$m': \qquad \frac{A\phi}{wh} = \beta(\alpha)V_{\alpha m}(m',h') \tag{2.29}$$

$$u: \qquad A = \beta(\alpha)V_{\alpha h}(m',h')f_u(u,h) \tag{2.30}$$

There is an important distinction between this model and the model employed by Aruoba, Waller and Wright (2008), where firms in the CM use physical capital instead of human capital. In this model, preference are quasilinear in hours of work as well, but production is carried out using effective hours of work. As will be clear below, this leads to different results in terms of trade determination and equilibrium. I note that even though first order conditions for x, m', h' depend on the current level of human capital (h), the joint distribution of (m,h) is still degenerate conditional on types. I assume that the initial distribution of knowledge (human capital) is degenerate across agents and there is no human capital depreciation when moving from the DM into the CM.[9]

[8] I am ignoring the illiquidity constraint for the moment since it turns out that in the steady state in does not bind.

[9] In this model, the only type of heterogeneity besides the discount factor differences is due to idiosyncratic trading shocks in the DM. The only variable it affects is the level of money holdings agents carry into the CM (buyers carry less or no cash, seller carry more cash and so on). Hence, conditional on types (patient versus impatient), when agents

Envelope conditions:

$$W_{\alpha m}(m,h) = \frac{A\phi}{wh} \tag{2.31}$$

$$W_{\alpha h}(m,h) = \frac{An}{h} + \frac{A}{f_u(u,h)}[1-\delta+f_h(u,h)] \tag{2.32}$$

There are two important points to notice from the envelope conditions which lead to different qualitative results: the first is that the marginal value of holding cash in the CM depends in the level of human capital, and the second is that the marginal value of human capital in the CM depends on the level of money balances. As we will see later on, the first will lead the DM terms of trade to depend on seller's and buyer's levels of human capital even though the buyer does not use human capital in the DM. This will also have different implications for the hold-up problem compared to Aruoba, Waller and Wright (2008).

Decentralized market

In the DM, buyers consume q amount of special goods and derive utility $u(q)$. A seller incurs a utility cost $c(q,\tilde{h})$ from producing q using labor and human capital.[10] As before, I consider single coincidence meeting, where the probability to meet a trading partner is σ. Let $V_\alpha(m,h)$ denote the value function of a type α agent entering the DM with m money holdings and h units of human capital. As will be shown from the

enter the CM they are identical in the level of human capital since we assume there is no depreciation when moving from DM into the CM.

[10]Here again as in the CM, production requires effective hours of work. Consider a production function $q=(n\tilde{h})^\gamma$ and disutility of labor in the DM measured by $v(n)=\frac{n^{1+\psi}}{1+\psi}$,where ψ is the Frisch labor supply elasticity. Then the utility cost of producing q for a seller with $\tilde{h}$ level of human capital is $c(q,\tilde{h})=\left(\frac{q^{\frac{1+\psi}{\gamma}}}{\tilde{h}^{1+\psi}}\right)\frac{1}{1+\psi}, \psi>0, \gamma\leqslant 1$.

Nash Bargaining stage, $q = q(m, h, \tilde{h})$ and $d = m$.

$$V_\alpha(m,h) = \sigma \int [-c(q(\tilde{m},\tilde{h},h),h) + W_\alpha(m + d(\tilde{m}),h)]dH(\tilde{m},\tilde{h})$$
$$+\sigma \int [u(q(m,h,\tilde{h})) + W_\alpha(m - d(m),h)dF(\tilde{h}) + (1-2\sigma)W_\alpha(m) \tag{2.33}$$

where each of the right hand side terms represents the expected value of being a seller, a buyer or a non-trader respectively. I now, turn to the terms of trade determination in the DM.

Terms of Trade Determination: Generalized Nash Bargaining

Terms of trade are determined by maximizing gains from trade:

$$\max_{q,d\leq m}[u(q) + W_\alpha(m-d,h) - W_\alpha(m,h)]^\theta[-c(q,\tilde{h}) + W_\alpha(\tilde{m}+d,\tilde{h}) - W_\alpha(\tilde{m},\tilde{h})]^{1-\theta}$$

where (m, h) are money holdings and human capital of the buyer and $(\tilde{m}, \tilde{h})$ are money holdings and human capital of the seller, and θ is the bargaining power of the buyer. Given the linearity in money holdings m of W_α, we have

$$W_\alpha(m+d,h) - W_\alpha(m,h) = d\frac{A\phi}{wh}$$

Hence the problem can be written as

$$\max_{q,d\leq m}[u(q) - d\frac{A\phi}{wh}]^\theta[-c(q,\tilde{h}) + d\frac{A\phi}{w\tilde{h}}]^{1-\theta}$$

The solution to the Nash Bargaining problem is as follows:

$$q(m,\tilde{m},h,\tilde{h}) = \begin{cases} q(m,h,\tilde{h}) & \text{if } m < m^*(h,\tilde{h}) \\ q^*(h,\tilde{h}) & \text{if } m \geq m^*(h,\tilde{h}) \end{cases}$$

$$d(m,\tilde{m},h,\tilde{h}) = \begin{cases} m & \text{if } m < m^*(h,\tilde{h}) \\ m^* & \text{if } m \geq m^*(h,\tilde{h}) \end{cases}$$

where q^*defines the optimal quantity produced during DM and can be found by solving

$$u'(q^*)h = \tilde{h}c_q(q^*,\tilde{h}) \tag{2.34}$$

and m^* defines the optimal amount of money required for purchasing q^*

$$m^* = \frac{w}{A\phi}\left[\theta\tilde{h}c(q^*,\tilde{h}) + (1-\theta)u(q^*)h\right] \tag{2.35}$$

For cases when $m < m^*, q(m,h,\tilde{h})$ is the q that solves $\frac{m\phi A}{w} = g(q,h,\tilde{h})$ with

$$g(q,h,\tilde{h}) = \frac{\theta u'(q)c(q,\tilde{h}) + (1-\theta)c_q(q,\tilde{h})u(q)}{\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}}$$

Implicitly, $q = q(m,h,\tilde{h})$, and $\partial q(m,h,\tilde{h})/m = \phi A/(wg_q) > 0, \partial q(m,h,\tilde{h})/\partial h = -g_h/g_q \leq 0, \partial q(m,h,\tilde{h})/\partial\tilde{h} = -g_{\tilde{h}}/g_q \geq 0$. Each partial derivative is signed as follows

$$g_q = \frac{c_q u_q(\frac{\theta u_q}{\tilde{h}} + (1-\theta)\frac{c_q}{h}) + \theta(1-\theta)(c_{qq}u_q - c_q u_{qq})(u-c)}{\left(\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}\right)^2} > 0$$

$$g_h = \frac{(1-\theta)c_q(\theta u_q c + (1-\theta)c_q u)}{h^2\left(\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}\right)^2} \geq 0$$

$$g_{\tilde{h}} = \frac{\theta u_q c_h \left(\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}\right) + \theta(1-\theta)u_q c_{qh}\left(\frac{u}{\tilde{h}} - \frac{c}{h}\right)}{\left(\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}\right)^2} + \frac{\frac{\theta u_q}{\tilde{h}^2}\left(\theta u_q c + (1-\theta)c_q u\right)}{\left(\frac{\theta u'(q)}{\tilde{h}} + \frac{(1-\theta)c_q(q,\tilde{h})}{h}\right)^2} \leq 0$$

It is important at this point to discuss whether in equilibrium, the amount of money held by each type of buyer satisfies $m < m^*$ or not. In the Lagos and Wright model as well as in Aruoba, Waller and Wright (2008), one can show that in equilibrium $m < m^*$and hence $d(m) = m$. This implies that agents hold less than the optimal amount of money, and all their money balances are used to purchase the DM good. In these papers this result is due to the lack of heterogeneity among sellers. In my model this is not the case. As one can see from equations (2.34) and (2.35), the optimal amount of money to purchase q^* will depend on seller's type. When deciding how much money to bring into the DM, buyers have to take into account the possible type of their future trading partner (seller's type). Meetings with less productive sellers (the ones having low human capital) will require more money balances and meetings with more productive sellers will require less money. Buyers know their own type but not the type of their trading partner. It might be very well the case that for some meetings $m < m^*$ and for some others $m > m^*$. This would introduce additional heterogeneity in the model and endanger its tractability. Whether one condition holds versus the other will depend on the structural parameters of the model. In appendix H, I solve a simple model with seller heterogeneity, and derive the condition under which $m < m^*$ is satisfied. Unfortunately, I cannot derive such a condition for the full model, due to its complexity. Hence, I solve the model for cases where $m < \min\{m_L^*, m_H^*\}$.

The marginal value of carrying money balances in the DM:

$$V_{\alpha m}(m,h) = \sigma \int W_{\alpha m}(m + d(\tilde{m}), h) dF(\tilde{m}) + \sigma \int [u'(q(m,h,\tilde{h})) q_m(m,h,\tilde{h})$$
$$+ W_{\alpha m}(m - d(m), h)] dF(\tilde{h}) + (1-2\sigma) W_{\alpha m}(m,h) \tag{2.36}$$

Substituting $W'_\alpha(m) = \frac{A\phi}{wh}$, and the Nash Bargaining outcomes $q_m(m,h,\tilde{h}) = \frac{\phi A}{w g_q}$, $d'(m) = 1$ above we have:

$$V_{\alpha m}(m,h) = (1-\sigma)\frac{A\phi}{wh} + \sigma \int [u'(q)\phi A/(w g_q(q,h,\tilde{h}) dF(\tilde{h}) \tag{2.37}$$

Marginal value of human capital:

$$V_{\alpha h}(m,h) = \sigma \int [-c_q(q(\tilde{m},\tilde{h},h),h) q_h(\tilde{m},\tilde{h},h) - c_h(q(\tilde{m},\tilde{h},h),h) + W_{\alpha h}(m + d(\tilde{m}), h)] dF(\tilde{m})$$
$$+ \sigma \int [u'(q(m,h,\tilde{h})) q_h(m,h,\tilde{h}) + W_{\alpha h}(m - d(m), h)] dF(\tilde{h}) + (1-2\sigma) W_{\alpha h}(m,h) \tag{2.38}$$

Using Envelope Conditions, partial derivatives from Nash Bargaining and some algebra, $V_{\alpha h}(m,h)$ becomes:

$$V_{\alpha h}(m,h) = \sigma \int [c_q(q,h)\tfrac{g_h(q,\tilde{h},h)}{g_q}) - c_h(q,h)] dF(\tilde{h}) - \sigma \int u'(q)\tfrac{g_h(q,h,\tilde{h})}{g_q(q,h,\tilde{h})} dF(\tilde{h})$$
$$+ \tfrac{A}{wh^2}[x - \pi - (z+\sigma)(\phi M - \phi m)] + \tfrac{A}{f_u(u,h)}(1 - \delta + f_h(u,h)) \tag{2.39}$$

Equilibrium Conditions

Substituting (2.37), (2.39), into (2.29) and (2.30) as well using $A\phi m/w = g(q,h,\tilde{h})$ we get the equilibrium conditions for a type α agent:

$$\frac{g(q,h,\tilde{h})}{mh} = \beta(\alpha)\frac{g(q',h',\tilde{h}')}{m'}\left[\frac{1-\sigma}{h'} + \sigma\int\frac{u'(q')}{g_q(q',h',\tilde{h}')}dF(\tilde{h})\right] \tag{2.40}$$

$$\frac{A}{f_u(u,h)} = \beta(\alpha)\left[\begin{array}{c} \sigma\int[c_q(q',h')\frac{g_h(q',\tilde{h}',h')}{g_q(q',h',\tilde{h}')}) - c_h(q',h')]dF(\tilde{h}) - \sigma\int u'(q')\frac{g_h(q',h',\tilde{h}')}{g_q(q',h',\tilde{h}')}dF(\tilde{h}) \\ \\ +\frac{A}{wh^2}[x' - \pi' - (z+\sigma)(\phi' M' - \phi' m')] + \frac{A}{f_u(u',h')}(1-\delta+f_h(u',h')) \end{array}\right] \tag{2.41}$$

$$U'(x) = \frac{A}{wh}$$

Given that the dynamics are computationally demanding, I choose to focus on the steady state analysis even though I am aware of potentially important dynamics of this model. At the steady state, I make use of the following: $M' = (1+z)M$, $m' = (1+z)m$, $\phi = (1+z)\phi'$, $w = \gamma[\int n_\alpha h_\alpha dG(\alpha)]^{\gamma-1}$ and profits, $\pi = (1-\gamma)[\int n_\alpha h_\alpha dG(\alpha)]^\gamma$. Note that the wage depends on aggregate effective labor supply in the economy. Hence, the two aggregate variables that enter each type's decision problem are M, $\int n_\alpha h_\alpha dG(\alpha)$, i.e. aggregate money supply and aggregate effective labor supply.

Steady state equations:

$$1+z = \beta(\alpha)\left[1-\sigma+\sigma\int\frac{u'(q_\alpha)h}{g_q(q_\alpha,h_\alpha,h_{\tilde{\alpha}})}dG(\tilde{\alpha})\right] \tag{2.42}$$

$$\frac{A}{f_u(u_\alpha,h_\alpha)}=\beta(\alpha)\left[\begin{array}{c}\sigma\int[c_q(q_\alpha,h_\alpha)\frac{g_h(q_\alpha,h_{\tilde{\alpha}},h_\alpha)}{g_q(q_\alpha,h_{\tilde{\alpha}},h_\alpha)})-c_h(q_\alpha,h_\alpha)]dG(\tilde{\alpha})-\sigma\int u'(q_\alpha)\frac{g_h(q_\alpha,h_\alpha,h_{\tilde{\alpha}})}{g_q(q_\alpha,h_\alpha,h_{\tilde{\alpha}})}dG(\tilde{\alpha})\\ +\frac{A}{\gamma\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}h^2}[x-(z+\sigma)(\phi M-\phi m_\alpha)-(1-\gamma)\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma}\\ +\frac{A}{f_u(u_\alpha,h_\alpha)}(1-\delta+f_h(u_\alpha,h_\alpha))\end{array}\right] \tag{2.43}$$

$$U'(x_\alpha)=\frac{A}{wh_\alpha} \tag{2.44}$$

$$n_\alpha=\frac{x_\alpha-z(\phi M-\phi m_\alpha)-\pi}{wh_\alpha} \tag{2.45}$$

From the Nash Bargaining solution we also have:

$$\frac{m_\alpha\phi A}{w}=g(q,h_\alpha,h_{\tilde{\alpha}})$$

As shown in the Nash Bargaining stage, the fact that the terms of trade now depend on seller's and buyer's human capital, results in novel insights regarding incentives to accumulate human capital. Equations (2.46) and (2.47) refer to the net surplus of a buyer and a seller entering the DM respectively.

$$NS^B=u(q)+W_\alpha(m-d,h)-W_\alpha(m,h)=u(q)-d\frac{A\phi}{wh} \tag{2.46}$$

$$NS^S=-c(q,\tilde{h})+W_\alpha(\tilde{m}+d,\tilde{h})-W_\alpha(\tilde{m},\tilde{h})=-c(q,\tilde{h})+d\frac{A\phi}{w\tilde{h}} \tag{2.47}$$

As one can see from equations (2.46), (2.47) the marginal value of holding money in the CM is inversely related to the level of human capital. That is, bringing an additional unit of money into the CM has a higher value for the less skilled agents. Bringing more human capital into the DM as a seller has two effects: it lowers the production cost, which tends to increase the net surplus, and it also reduces the marginal value of carrying cash into the CM, which lowers the net surplus. Bringing more human capital into DM as a buyer lowers the marginal cost of carrying no cash into the CM, so it increases the net surplus of the buyer. For cash constrained buyers, this implies that the higher the level of human capital of the buyer, the lower the amount of production q in the DM $(\partial q(m, h, \tilde{h})/\partial h < 0)$. When the buyer has full bargaining power $\theta = 1$, to capture all the benefits of the surplus s/he creates by bringing m to the DM, $\partial q(m, h, \tilde{h})/\partial h = 0$. The quantity traded in the DM is insensitive to the buyer's level of human capital. Giving full bargaining power to the buyer implies that the incentives to invest in human capital will be determined by the returns in the CM as well as its effect on the seller's production in the DM. From a seller's perspective, a higher $\tilde{h}$ in the DM has two opposing effects: it lowers the cost of production when trade occurs, but any amount of m received in exchange is associated with a lower marginal value in the CM. The net effect on the quantity traded q is that more skilled sellers lead to higher q, $\partial q(m, h, \tilde{h})/\partial \tilde{h} > 0$.

2.5 Numerical Solution of the Second Model

In this subsection, I show simulations to check the long run effect of inflation on the real economy, the distribution of wealth, and welfare. Given that the wage is set at the marginal product of aggregate effective labor supplied in the economy, there are potential externalities taking effect through wages, so I experiment with the returns to scale technology in the CM[11]. When I refer to comparisons between an heterogenous agents model and a model with a degenerate distribution the baseline experiment underlined is as follows: I first compute the steady state results under the assumption that all agents have the same discount factor, which leads to a degenerate distribution of assets. I then introduce a group of agents with a lower discount factor and track down how the behavior of the old group changed.

Parametrization

$$\text{CM} \ : \ Y_c = ZF(L) = ZL^{\gamma}, L \ = \int n_{\alpha} h_{\alpha} dG(\alpha)$$

$$f(u,h) = uh^{\mu}, \ \mu \in (0,1]$$

$$U(x) = B\log(x)$$

$$\text{DM} : u(q) = \frac{(q+b)^{1-\eta} - b^{1-\eta}}{1-\eta}$$

$$c(q,\tilde{h}) = \left(\frac{q^{\frac{1+\psi}{\gamma}}}{\tilde{h}^{1+\psi}} \right) \frac{1}{1+\psi}$$

[11]Experimenting with returns to scale in the DM is of little relevance, since production takes place in single-agent firms (simple individual production) and there no channel in the production technology connecting different types of agents.

Table 2.3: Benchmark Parameter Values - Human Capital Augmented Model

Parameter	**Value**	**Description**
η	1	coefficient of risk aversion
θ	0.745	buyer's bargaining power
σ	0.26	probability of bilateral meetings
B	1.30	constant
b	0.0001	constant
ψ	1	labor elasticity
γ	0.85	returns to scale
ω_L	0.5	share of low-type (impatients
Z	0.1985	constant
μ	0.5	returns to human capital accumulation
δ	0.04	human capital depreciation
A	4	disutility of labor in CM
β_L	0.94	discount factor for low-type
β_H	0.9615	discount factor for high-type

The cost function $c(.)$ comes from the production technology $q = (n\tilde{h})^\gamma$ and the disutility of labor in the DM measured by $v(n) = \frac{n^{1+\psi}}{1+\psi}$,where ψ is the Frisch labor supply elasticity[12]. Then the utility cost of producing q for a seller with $\tilde{h}$ level of human capital is $c(q,\tilde{h}) = \left(\frac{q^{\frac{1+\psi}{\gamma}}}{\tilde{h}^{1+\psi}}\right)\frac{1}{1+\psi}, \psi > 0, \gamma \leqslant 1$.Table 2.3 summarizes the benchmark parametrization used in the model.

[12]Here, as assumed in the CM, production is carried out using effective units of labor.

Time period in the model is one year. I use the values from Aruoba, Waller and Wright (2008) for the overlapping parameters and functional forms. Human capital is increased via a concave function, following (Ortigueira, 2000).

Constant returns to scale and homogenous agents

I first show what the long run effects of a higher money growth rate are on hours of work, time devoted to human capital accumulation and consumption, when there is no ex-ante heterogeneity. Agents can fully smooth their idiosyncratic trade shocks by adjusting their non-leisure time in the CM, which leads to a degenerate distribution of assets. As we can see from Figure 2.2, a higher rate of money growth leads to less time spent in education, a lower steady state level of human capital, lower consumption, and a constant amount of hours devoted to market activity. For the working hours to remain constant it must be the case, as one can see from equation (2.45), that CM consumption and human capital are decreasing by exactly the same amount. In order to understand the intuition behind the human capital response to inflation in the steady state, one should make clear where the returns to accumulating human capital come from. There are benefits to bringing human capital as a seller into the DM since it lowers production cost, so anything that taxes DM activity (in this case inflation) will lower the returns to human capital in the DM. Returns to holding human capital in the CM come in the form of higher labor income. Hence, in the face of a higher money growth rate, there are two opposing incentives to accumulation of human capital. Here, accumulation of human capital comes in the form of a greater share of time devoted to skill acquisition. From the law of motion for human capital, one can derive the steady state level of time devoted to education as follows:

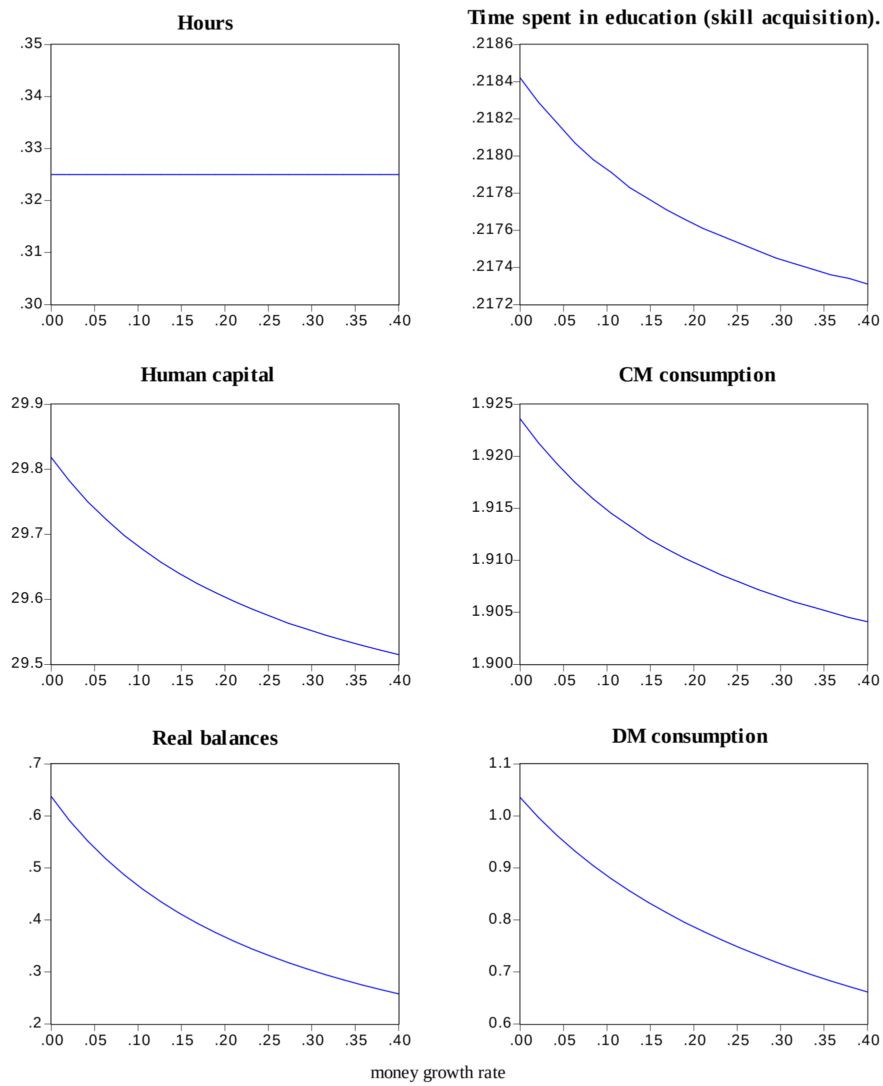
$$u_\alpha = \delta h_\alpha^{1-\mu}$$

Hence, agents with a higher level of human capital devote more time to education, but do so at a decreasing rate. Marginal utility of CM consumption is inversely related to the level of human capital (equation (2.44)).This means that as human capital decreases, marginal utility of consumption increases, which indicates a falling level of consumption. The main result of the model, when I abstract from heterogeneity, is that monetary policy has real effects on aggregate consumption, human capital accumulation, and time devoted to skill acquisition, and has no effect on aggregate hours of work. It must be noted that the effect of monetary policy is very small for CM variables and has a much bigger effect on DM consumption.

Constant returns to scale and heterogeneous agents

I now introduce a new group of agents with a lower discount factor ("the impatients"). As previously stated, more patient agents accumulate more money and more human capital. The introduction of heterogeneity in this environment does not change any of the qualitative results in the degenerate case, except for the hours of work. As one can see from Figure 2.3The high types, which happen to be the rich group, tend to work more in the face of higher inflation. It is important to note that this does not

Figure 2.2: The long-run effects of expansionary monetary policy - Costant Returns to Scale technology



mean that the rich group is experiencing an increase in welfare cost relative to the poor, as will be shown in the welfare analysis section. CM variables are little affected by changes in money growth rate, as before. DM trade, on the other hand is more influenced by these changes. The quantity of goods being produced in the DM is determined by the type of buyer and seller that meet in the pairwise meetings. The largest amount of DM trade occurs between a high-type buyer and high-type seller.

Given the redistributive effects of inflation, we are also interested in how inflation impacts wealth and human capital distribution in the long run. Figure 2.4 shows that dispersion of money holdings and human capital increases with inflation. This implies that, in steady state, real balances and human capital decline at a faster pace for low types as money growth rate increases.

Figure 2.5 shows the impact of changes in the money growth rate on aggregate variables in the economy. Monetary policy has real effects. Even though these effects are quite small in terms of quantities, they can be significant in welfare terms, as will be shown in the next section.

Decreasing Returns to Scale Technology

Besides the lump sum transfer of money in the amount $\tau = zM$ that agents get in the CM, returns to scale technology is the only other channel through which another aggregate (total supply of labor) can interact with individual variables, and potentially lead to qualitatively different results. As mentioned, before given that wage is

Figure 2.3: Type-specific long-run effects of expansionary monetary policy - Constant returns to scale

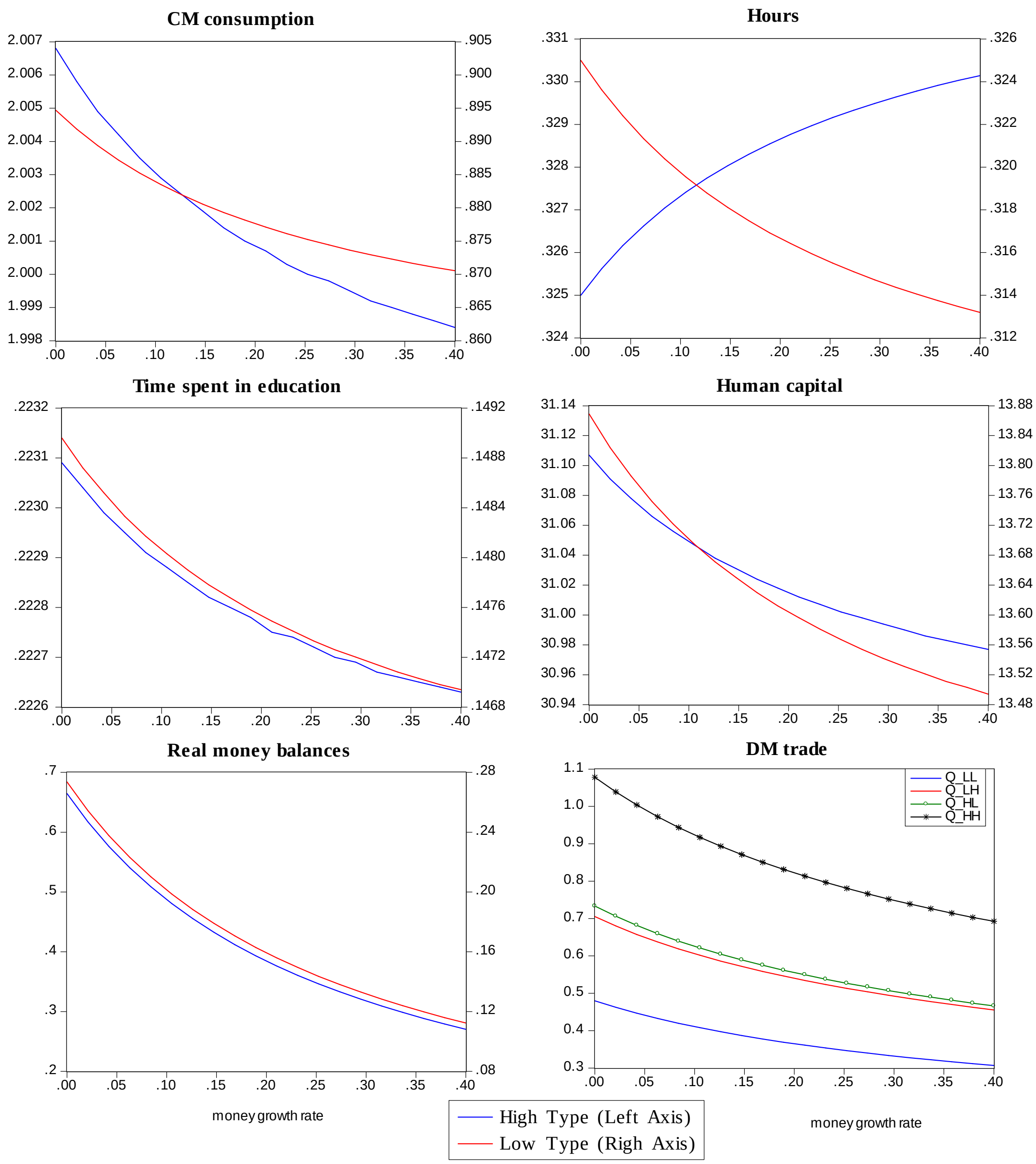
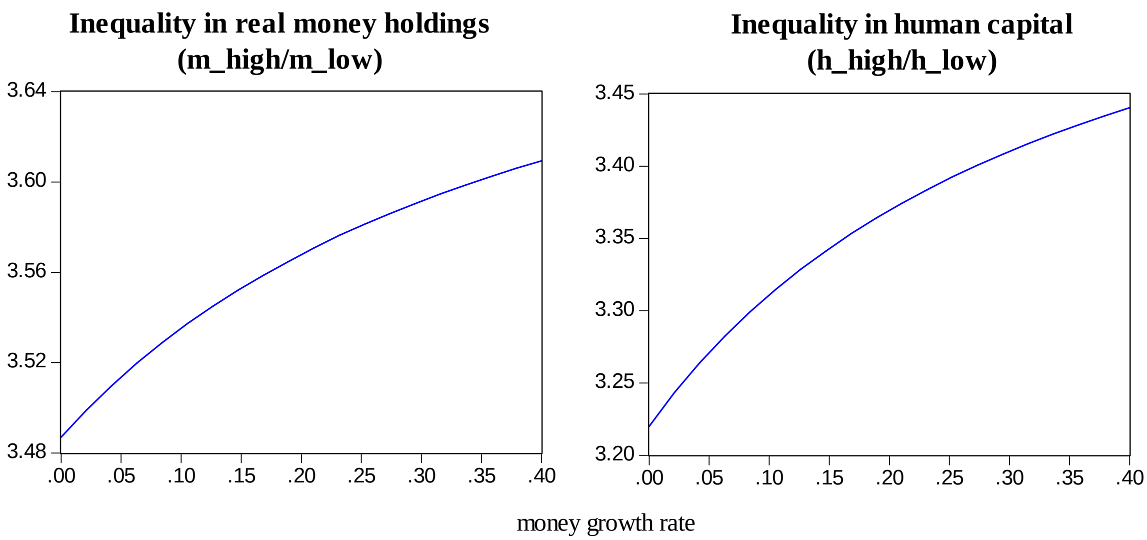
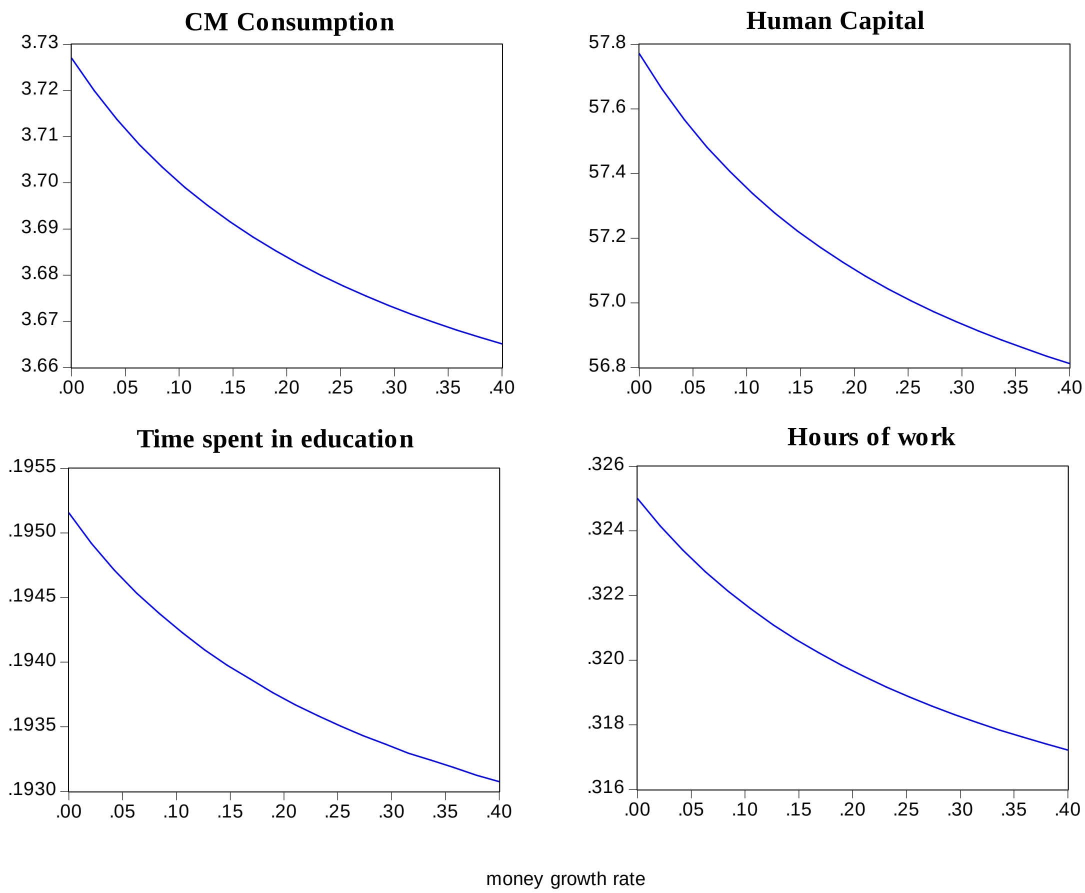


Figure 2.4: Distribution of wealth and human capital - Constant returns to scale



set at the marginal product of aggregate effective labor supplied in the economy, there are potential externalities taking effect through wages. I note that there are different ways as well to allow for externalities in this model. For example, human capital production function may depend on the economy wide level of education. To save space, I do not present the degenerate case results for decreasing returns to scale, but they are qualitatively the same as Figure 2.2. Figure 2.6 shows the effect of monetary expansion on hours of work, education, and human capital for each type of agent under DRTS technology. Introducing DRTS changes steady-state results in a substantial way. First, there is a change in roles. The impatient agents in this case are the

Figure 2.5: Long run effect of inflation on aggregate variables - Constant returns to scale



ones who hold more money, accumulate more human capital, work and study more. DRTS makes the wage sensitive to the heterogeneity. Under DRTS, the introduction of a different group of agents, which will ultimately supply different levels of effective labor in the CM, will affect the wage rate, and hence the CM margin of human capital accumulation. It is the case that as the money growth rate increases, DRTS makes the return to accumulating human capital in the CM dominate the negative impact on the DM return for patient (poor) agents. Second, due to the additional redistributive effects provided, now, by a non-constant wage rate, which depends on the economy's aggregate effective labor, CM consumption is decreasing for the low type. Unlike in the CRTS case, steady state human capital and time spent in skill acquisition are increasing in the rate of money growth.[13]

Figure 2.7 shows that the relatively richer agents (impatients) react to changes in the money growth rate by reducing real balances faster than the poor agents do. The steady state level of human capital, on the other hand, is increasing at a slower pace for the richer agents. This implies that in the long run differences in wealth and human capital are diminishing.

Figure 2.8 displays the long run effect of changes in money growth rate for aggregate variables. Aggregate human capital, CM consumption, and time spent in skill acquisition increase with the money growth rate, whereas aggregate hours of work decrease.

[13]It must be noted that such overturning results do not hold for all $\gamma < 1$. The degree of decreasing returns to scale must be low enough to yield this sections results. In my numerical examples, $\gamma \leq 0.85$.

Figure 2.6: Type specific long-run effects of monetary expansion. Decreasing Returns to Scale CM production technology

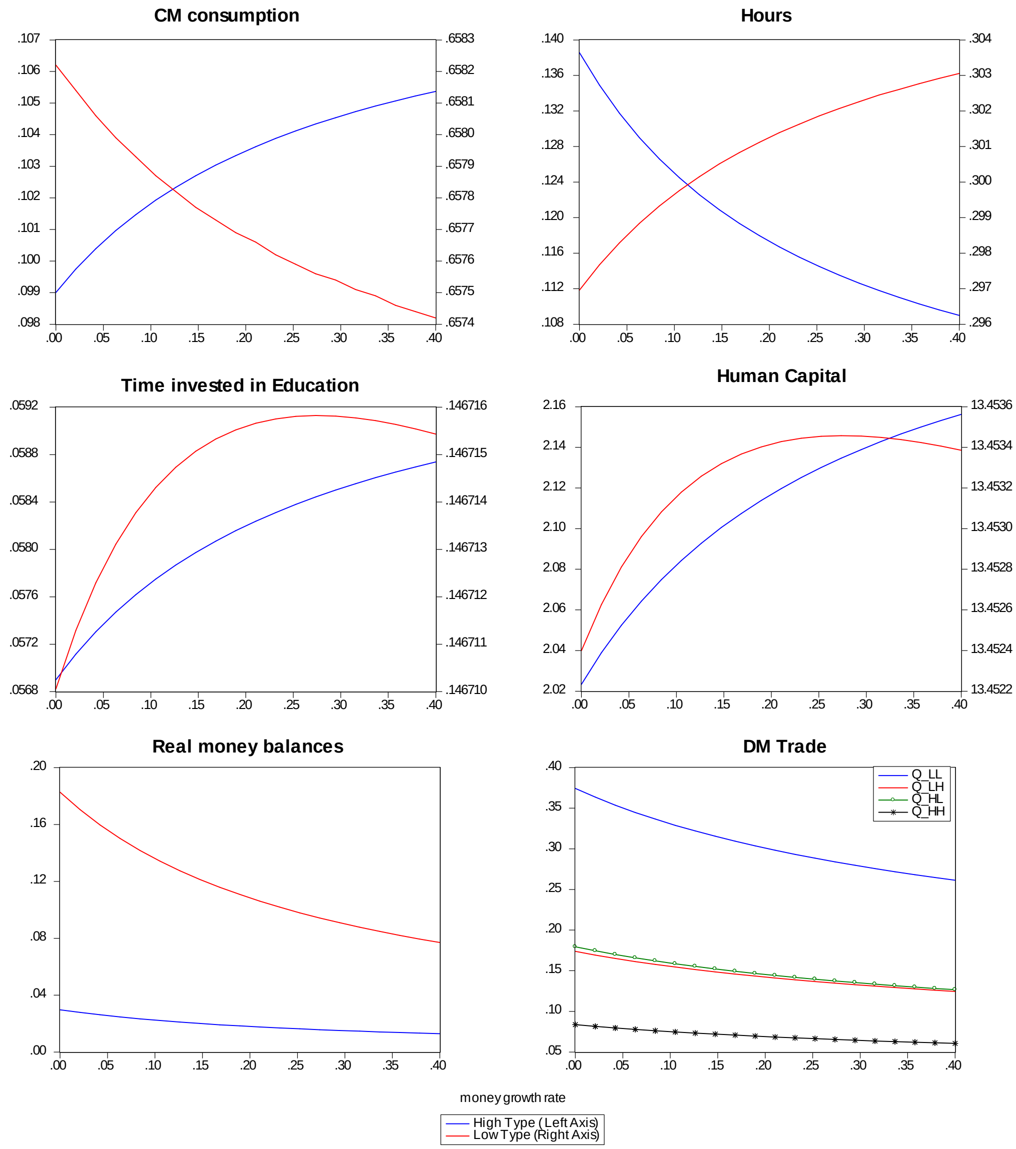
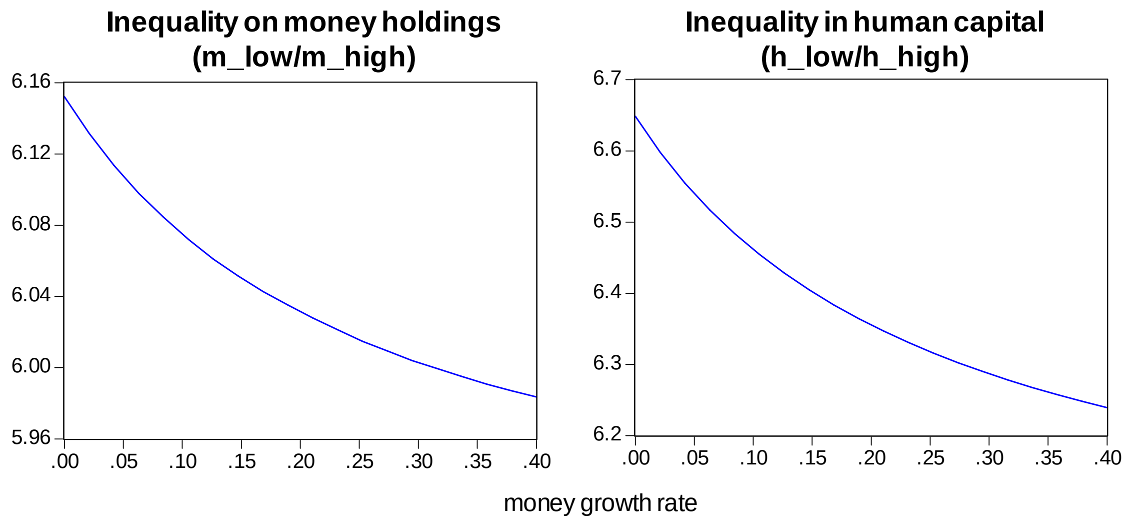


Figure 2.7: Dispersion in wealth and human capital - Decreasing Returns to Scale CM production technology



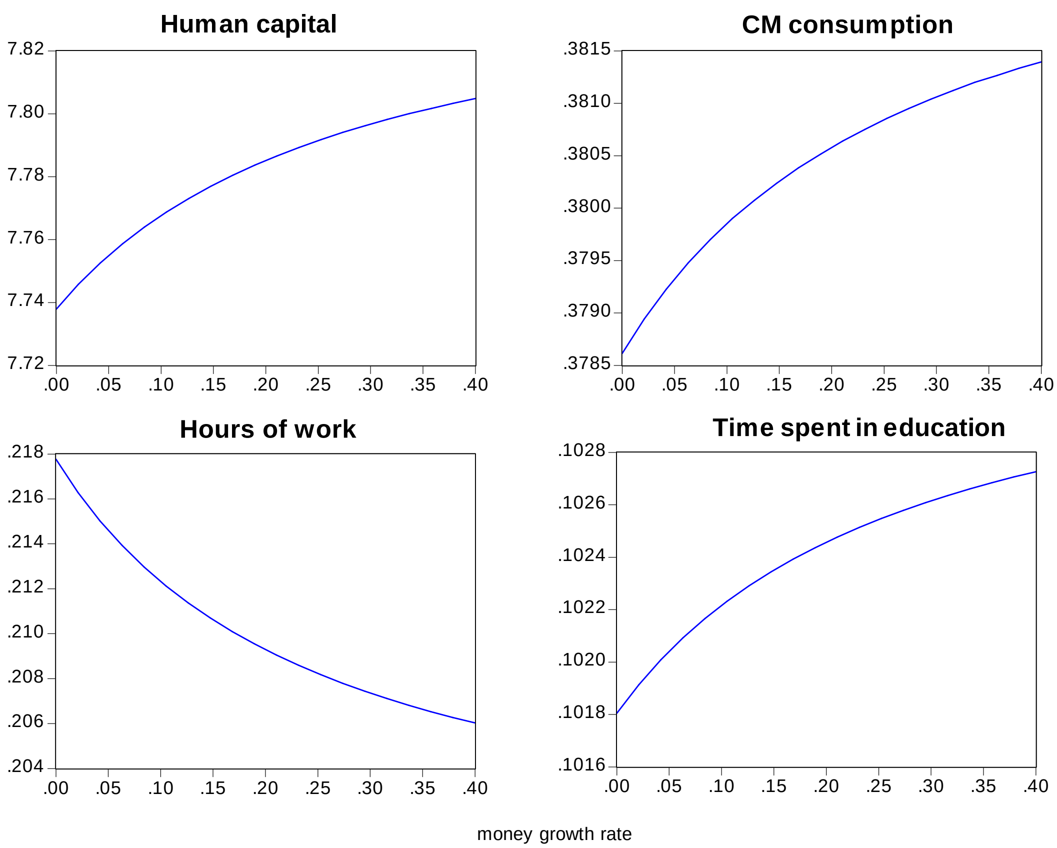
Discussion on Decreasing Returns to Scale

In this section I address the results obtained under the DRTS calibration, with particular focus on the conditions under which, the patient agents in the steady state accumulate *less* human capital. The first order condition that determines the equilibrium amount of human capital (hours spent in education) is given by equation (2.30), which I repeat here for convenience.

$$A = \beta(\alpha)V_{\alpha h}(m', h')f_u(u, h)$$

This equation implies that agents with a higher discount factor (the patient) require a lower future marginal value of human capital (lower $V_{\alpha h}(m', h')$). The concavity of

Figure 2.8: The long run effect of monetary policy on aggregate variables - Decreasing returns to scale



the value function, implies a tendency for the patient to accumulate more human capital. Another additional factor however, is the nature of the human capital production function $f(u,h)$. Heckman (1975) and Heckman, Lochner and Taber (1998) estimates show that human capital production function displays strong increasing returns to scale. As it is the case in this paper as well, the human capital production function exhibits increasing returns to scale.[14] Equation (2.30) reconciles the human capital production function and the discount factor. For an increasing returns to scale human capital production function, the more patient agents have an incentive to invest *less* time in education and accumulate *less* human capital today since the return to education will be increasing in the level of human capital.[15] Hence, for a high discount factor there are two opposing tendencies in the accumulation of human capital and time invested in education. In order to see how the returns to scale in human capital production interact with the returns to scale in the CM good production, I rewrite

[14] It must be noted that the results are sensitive to the degree of IRTS for $f(u,h)$.

[15] Consider a general human capital production function widely used in the literature $f(u,h) = (uh)^{\mu}$. For DRTS, $\mu < 1/2$, for CRTS $\mu = 1/2$ and for IRTS $\mu > 1/2$. This functional form implies that at the steady state $f_u(u,h) = \mu\delta^{\frac{\mu-1}{\mu}} h^{2-\frac{1}{\mu}}$. This implies that

$$f_{uh}(u,f) = \mu\delta^{\frac{\mu-1}{\mu}}(2-\tfrac{1}{\mu})h^{1-\frac{1}{\mu}} \begin{cases} > 0, \text{ if IRTS, } \mu > 1/2 \\ < 0, \text{ if DRTS, } \mu < 1/2 \\ = 0, \text{ if CRTS, } \mu = 1/2 \end{cases}$$

the steady state equation (2.43) associated with equation (2.30).

$$A = \beta(\alpha) \left[\begin{array}{c} \sigma \displaystyle\int [c_q(q_\alpha, h_\alpha) \tfrac{g_h(q_\alpha, h_{\tilde{\alpha}}, h_\alpha)}{g_q(q_\alpha, h_{\tilde{\alpha}}, h_\alpha)}) - c_h(q_\alpha, h_\alpha)] dG(\tilde{\alpha}) - \sigma \int u'(q_\alpha) \tfrac{g_h(q_\alpha, h_\alpha, h_{\tilde{\alpha}})}{g_q(q_\alpha, h_\alpha, h_{\tilde{\alpha}})} dG(\tilde{\alpha}) \\ \\ + \tfrac{A}{wh^2} [x - (z+\sigma)(\phi M - \phi m_\alpha) - \pi] \\ \\ + \tfrac{A}{f_u(u_\alpha, h_\alpha)} (1 - \delta + f_h(u_\alpha, h_\alpha)) \end{array} \right] f_u(.) \tag{2.48}$$

The first line on the right hand side reflects the returns to human capital in terms of DM production, taking into consideration the effect it has on the terms of trade. The second and the third line reflect the expected returns to human capital when entering the CM as a buyer, a seller or a non-trader.[16] Ignoring for a moment the DM aspect of human capital accumulation, a higher discount factor is associated on one hand, by a higher human capital stock (look at the first term in the second line), and on the other hand, with a lower human capital stock if there are increasing returns to human capital production function, and higher (constant) human capital stock in the case of decreasing (constant) returns to scale in the human capital production technology. In this chapter, I use an increasing returns to scale technology. This implies that for the patient agent there are two opposing incentives in terms of accumulating human capital.

Under CRTS technology in the production of the CM good, $w = 1$ and $\pi = 0$, the tendency to accumulate higher levels of human capital for the patient agent

[16] Expected returns, depending on the various trading partner matches.

dominates the increasing returns to scale factor in human capital production. On the other hand, decreasing returns to scale technology for CM production lead to $w = \gamma \left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}$ and $\pi = (1-\gamma)\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma}$, which implies that another factor, namely the wage rate can alter the incentives to accumulate human capital. Equation (2.48) becomes

$$A = \beta(\alpha)\left[\begin{array}{c} \sigma \displaystyle\int [c_q(q_\alpha, h_\alpha)\tfrac{g_h(q_\alpha,h_{\tilde{\alpha}},h_\alpha)}{g_q(q_\alpha,h_{\tilde{\alpha}},h_\alpha)}) - c_h(q_\alpha,h_\alpha)]dG(\tilde{\alpha}) - \sigma \int u'(q_\alpha)\tfrac{g_h(q_\alpha,h_\alpha,h_{\tilde{\alpha}})}{g_q(q_\alpha,h_\alpha,h_{\tilde{\alpha}})}dG(\tilde{\alpha}) \\ + \dfrac{A}{\underbrace{\gamma\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}}_{w} h^2}[x-(z+\sigma)(\phi M - \phi m_\alpha) - \underbrace{(1-\gamma)\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma}}_{\pi} \\ + \frac{A}{f_u(u_\alpha,h_\alpha)}(1-\delta+f_h(u_\alpha,h_\alpha)) \end{array}\right] f_u(.)$$

A patient agent (high discount factor), through the wage-channel has the incentive to reduce the amount of human capital stock, since higher human capital depresses the wage rate. In this case for strong enough decreasing returns in the CM production technology and increasing returns technology in human capital production, can lead to patient agents to accumulate less human capital, as shown in the numerical results in this section. I note that such a result is sensitive to the returns to scale parameters in both production technologies and hence it is *not* a general result

2.6 Welfare Analysis

As mentioned previously, looking at the behavior of hours of work in the above graphs can be misleading in terms of welfare analysis. I measure welfare as expected utility in the steady state (2.49), which consists of expected DM utility and expected CM utility. Similar to the previous section, in order to assess the welfare effects of inflation, I look at the steady state expected utility of an agent type α entering the DM with (m,h)

$$(1-\beta(\alpha))V_\alpha = \sigma E[u(q_\alpha)-c(q_\alpha,h)] + U(x_\alpha) - E(n_\alpha+u_\alpha) \tag{2.49}$$

The first term denotes the expected utility from the DM. Note that, here, I take the expectation not only with respect to the trading status, but also with respect to all four possible meetings (type α_i buyer meets type α_j seller, $(i,j)\in(L,H)$). I compute the expected working hours as I did in the previous section:

$$E(n_\alpha+u_\alpha) = \bar{n}_\alpha + u_\alpha = u_\alpha + \frac{x_\alpha - \phi(\sigma+z)(M-m_\alpha)-\pi}{wh_\alpha} \tag{2.50}$$

We are interested in how higher rates of money growth rate affect welfare :

$$\frac{\partial(1-\beta(\alpha))V_\alpha}{\partial z} = \sigma\frac{\partial E[u(q_\alpha)-c(q_\alpha,h_\alpha)]}{\partial z} + U'(x_\alpha)\frac{\partial x_\alpha}{\partial z} - \frac{\partial(\bar{n}_\alpha+u_\alpha)}{\partial z} \tag{2.51}$$

$$\frac{\partial(\bar{n}_\alpha+\bar{u}_\alpha)}{\partial z} = \frac{\partial u_\alpha}{\partial z} + \frac{\left\{\begin{array}{c}\left[-\phi(M-m_\alpha)-(\sigma+z)(\frac{\partial\phi M}{\partial z}-\frac{\partial\phi m_\alpha}{\partial z})+\frac{\partial x_\alpha}{\partial z}-\frac{\partial\pi}{\partial z}\right]wh_\alpha \\ \\ -[x_\alpha-\phi(\sigma+z)(M-m_\alpha)-\pi][\frac{\partial w}{\partial z}h_\alpha + w\frac{\partial h_\alpha}{\partial z}]\end{array}\right\}}{w^2h_\alpha^2} \tag{2.52}$$

Substituting $w = \gamma \left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}$ above

$$\begin{aligned}\frac{\partial(\bar{n}_\alpha+\bar{u}_\alpha)}{\partial z} &= \frac{\partial u_\alpha}{\partial z} + \frac{\left[-\phi(M-m_\alpha)-(\sigma+z)(\frac{\partial \phi M}{\partial z}-\frac{\partial \phi m_\alpha}{\partial z})\right]}{\gamma\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1} h_\alpha} + \frac{\frac{\partial x_\alpha}{\partial z}-\frac{\partial \pi}{\partial z}}{\gamma\left[\int [n_\alpha h_\alpha dG(\alpha)^{\gamma-1}\right] h_\alpha} \\ &- \frac{[x_\alpha-\phi(\sigma+z)(M-m_\alpha)-\pi][\frac{\partial w}{\partial z}h_\alpha+\gamma\left[\int n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}\frac{\partial h_\alpha}{\partial z}]}{\gamma\left[\left[\int [n_\alpha h_\alpha dG(\alpha)\right]^{\gamma-1}\right]^2 h_\alpha^2}\end{aligned} \tag{2.53}$$

What is important to notice is that, now, there is another channel, namely the human capital channel, through which a higher money growth rate can redistribute among types. Given that it would be very tedious to show analytically how each welfare term is affected as we change the rate of money growth, and how it differs across agents I choose to rely on the numerical results. I proceed by examining each of the three terms that affect welfare: *expected utility from the DM , utility from CM consumption, and expected utility from leisure in the CM.*

In Figure 2.9, I decompose CM expected utility into utility from leisure and utility from consumption. The poor agents (patients) benefit from the introduction of the impatients because, as compared to the degenerate distribution case, this leads to an increase in CM utility and welfare from inflation. DM welfare, on the other hand, is decreasing for both types.

For completeness, I do the same exercise with CRTS technology. Figure 2.10 shows that the presence of a different group of agents in the economy leads to different welfare costs of inflation. The relatively poor agents bear a higher cost of inflation as compared to the richer agents.[17]

[17] The quatification of such a statement can be found in Table 4.

Figure 2.9: Welfare effects on inflation - Decreasing Returns to Scale

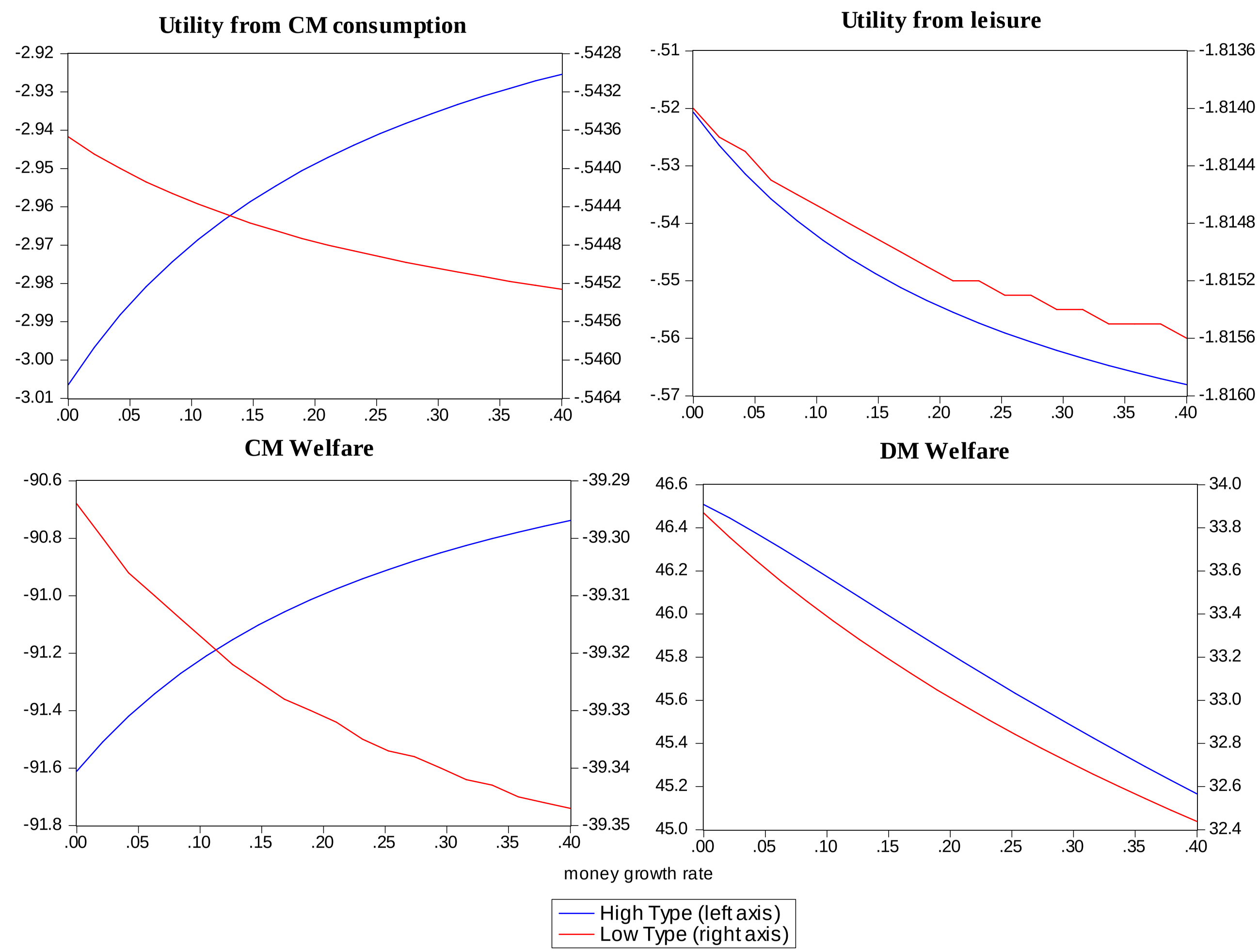
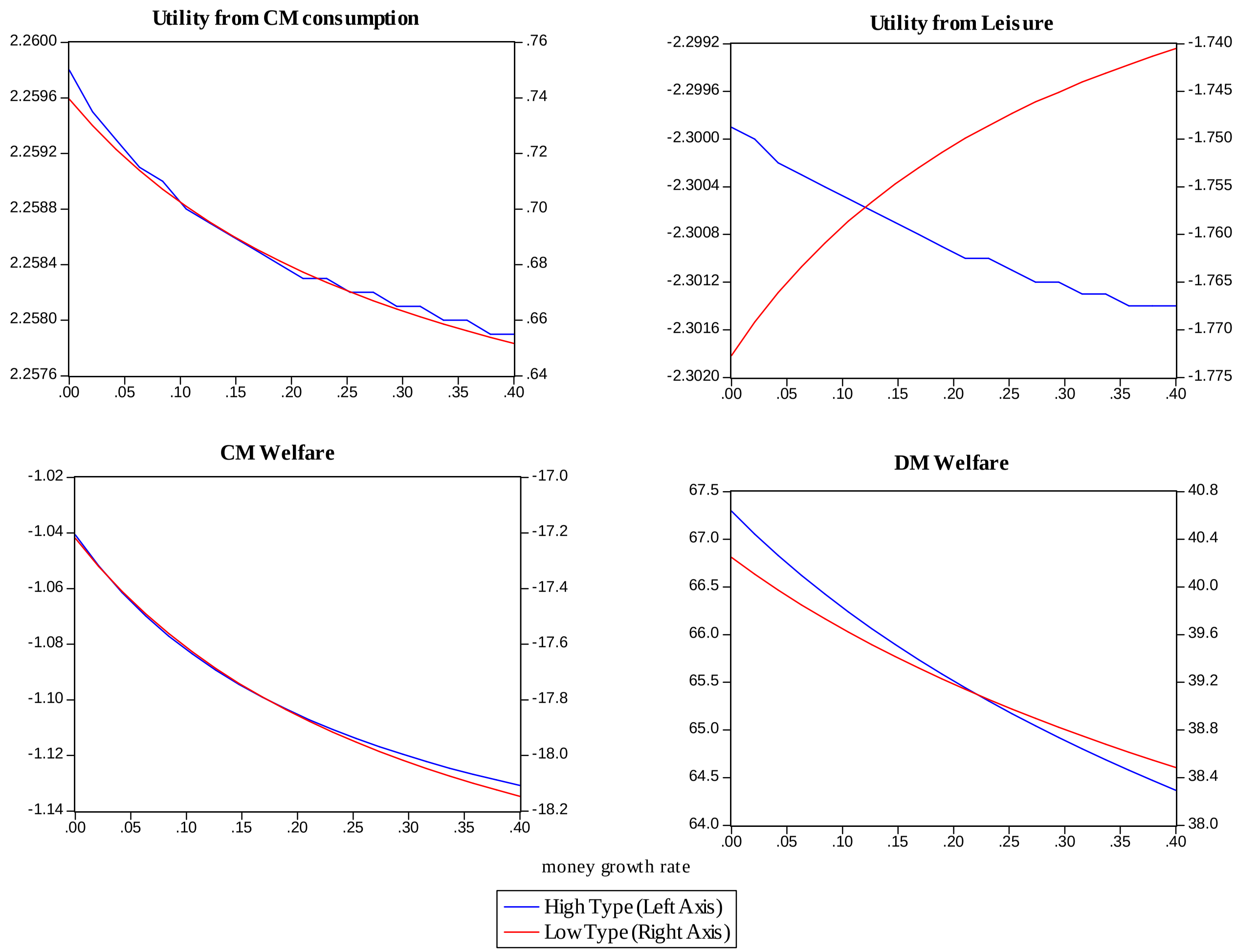


Figure 2.10: Welfare consequences of inflation - Constant Returns to Scale



In Tables 2.4 and 2.5, I present the welfare cost of moving from 0% to 10% inflation rate for various parameter values. For CRTS technology (Table 2.5)in CM production I observe the following: for all parameter values, poor agents suffer more from inflation, indicating a regressive inflation tax. The lower the weight of the poor agents, and the bigger the difference in discount factors, the greater is the dispersion in welfare cost between different agents. Allowing for DRTS technology (Table 2.4), on the other hand, offers another channel of redistributive effects. In this case, the relatively poor agents (high types in this case) benefit from inflation, and the rich agents bear the cost of higher money growth rates. The welfare cost for rich agents is systematically smaller than in an economy with CRTS technology (constant wages). Overall, allowing for another productive asset, whose return depends on economy-wide prices, allows for yet another channel of redistribution through which inflationary policies operate. When that is the case, considering representative agent models overstates the welfare cost of inflation as well as hiding the fact that a section of the economy can actually benefit from it.

2.7 Conclusion

This chapter studies the long-run redistributive effects of monetary policy in a micro-founded model of money. It builds on the search-theoretic model of Lagos and Wright (2005) in two important dimensions. First, I introduce heterogeneity while keeping the distribution of money holdings tractable. This version of the model allows us to

Table 2.4: Welfare cost of moving from 0% to 10% inflation - Decreasing Returns to Scale

(% consumption)

Parameter	**High Type**	**Low Type**	**Deg. Dist. ($\beta = \beta_H$)**
Decreasing Returns to Scale			
Case 1 - Benchmark Calibration (Table 2.3)	-0.94	1.69	2.27
Case 2 - Case 1 with $\psi = 0.2$	0.30	2.96	4.00
Case 3 - Case 1 with $\psi = 0.5$	-0.24	2.33	3.11
Case 4 - Case 1 with $\omega_L = 0.3$ (lower share of the impatients)	-1.36	1.54	2.27
Case 5 - Case 1 with $\mu = 0.5$	-1.00	1.72	2.20
Case 6 - Case 1 with $\mu = 0.7$	-0.89	1.63	2.36
Case 7 - Case 1 with $\beta_L = 0.94$	-0.06	1.96	2.27
Case 8 - Case 7 with $\omega_L = 0.35$	-0.34	1.87	2.27
Case 9 - Case 1 with $\theta = 0.5$	-0.15	1.59	2.10
Case 10 - Case 1 with $\gamma = 0.9$	-2.07	1.71	2.42
Case 11 - Case 10 with $\beta_L = 0.92$	-1.59	1.87	2.35

Table 2.5: Welfare cost of moving from 0% to 10% inflation - Constant Returns to Scale

(% consumption)

Parameter	**High Type**	**Low Type**	**Deg. Dist.** ($\beta = \beta_H$)
Constant Returns to Scale			
Case 1 - Benchmark parameters with $\gamma = 1$	2.63	3.89	2.72
Case 2 - Case 1 with $\omega_L = 0.7$ (higher share of "impatients")	2.55	3.14	2.72
Case 3 - Case 1 with $\mu = 0.5$	2.58	2.88	2.63
Case 4 - Case 1 with $\omega_L = 0.3$	2.75	6.71	2.72
Case 4 - Case 1 with $\theta = 1$	4.05	7.07	4.58
Case 5 - Case 1 with $\theta = 0.3$	2.39	2.80	2.41
Case 6 -Case 1 with $\theta = 0.5$	2.46	3.14	2.50
Case 7 - Case 1 with $\beta_L = 0.92$	2.65	4.21	2.63

examine the redistributive effects of changes in the money growth rate when money is injected via lump-sum transfers. Heterogeneity in discount factors results in a regressive inflation tax. Wealthy agents are less affected by the inflation tax than the less wealthy.

Second, I introduce human capital as a productive asset, which can be used in both DM and CM markets. This capital provides a link between the inflation tax in the DM and CM activity, breaking in this way the super-neutrality of money in the CM. I examine the effect of an increase in the money growth rate on output, welfare and the distribution of wealth and human capital accumulation. I discover two channels of redistributive effects of inflation. One is the usual effect generated by lump-sum transfers of money injected into the economy. The other effect is through the wage rate, which under decreasing returns to scale technology depends on economy-wide effective labor. My numerical results show that inflationary monetary policy can lead to a long-run increase in output, consumption, and time spent in skill acquisition activities and a decrease in the time spent working as well as a lower dispersion in the distribution of wealth and human capital.

Appendix A

Endogeneizing Information Processing Capacity (κ)

The fixed point solution for the *linear cost* case to the problem under endogeneous κ is the one that solves the following system of equations:

$$2^{2\kappa^A} = (1-\varphi)\left(\frac{\sigma_a}{\sigma_{a_i}}\right)2^{\kappa}$$

$$\mid \pi_{33} \mid \left(\frac{1}{1-\delta}\right)^2 \sigma_{a_i}^2 2^{2(\kappa^A-\kappa)} = 2^{\kappa}$$

and

$$\frac{\varphi-\gamma}{\psi+\gamma\delta}a_t = \frac{1-\varphi}{1-\delta}\left(1-2^{-2\kappa^A}\right)$$

The φ which verifies the inital guess is the solution to the following equation

$$K_1\varphi^2 + K_2\varphi + K_3 = 0$$

where $K_1 = -\left(1-\delta+\frac{c}{|\pi_{33}|\ln(2)\sigma_a^2}\left(\psi+\gamma\delta\right)\right)$, $K_2 = 1-\delta+\gamma-\gamma\delta+2\frac{c}{|\pi_{33}|\ln(2)\sigma_a^2}\left(\psi+\gamma\delta\right)$, $K_3 = \gamma(1-\delta)+\frac{c}{|\pi_{33}|\ln(2)\sigma_a^2}\left(\psi+\gamma\delta\right)\left(1-(1-\delta)^2\right)$

Appendix B

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.

Appendix C

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 I have:

$$\bar{C}^{-\gamma} w = \theta \bar{L} \tag{C.1}$$

$$1 = \beta(1 + r - d) \tag{C.2}$$

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

$$w = \delta \bar{K}^{\alpha} \bar{L}^{\delta - 1} \tag{C.3}$$

$$r = \alpha \bar{K}^{\alpha - 1} \bar{L}^{\delta} \tag{C.4}$$

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

$$\bar{C} = \bar{Y} + d\bar{K} \tag{C.5}$$

$$\bar{Y} = \bar{K}^{\alpha} \bar{L}^{\delta} \tag{C.6}$$

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

Appendix D

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 $z_t = 1, L_t = \bar{L}$ I 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 I 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,

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

Appendix E

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.

E.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.

E.2 Information rate of discrete parameter multi-dimensional Gaussian processes

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

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, 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 chapter 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[1]. 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.

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

$l_t^{FA} = D(L)\varepsilon_t$, $k_t^{FA} = E(L)\varepsilon_t$, $l_{it}^{*A} = M^L(L)\varepsilon_t + N^L(L)\eta_{it}^L$, $k_{it}^{*A} = M^K(L)\varepsilon_t + N^K(L)\eta_{it}^K$, 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 chapter 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 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 $u_{it} \sim WN(0, \sigma_u^2), \psi_{it} \sim WN(0, 1)$.

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

Appendix F

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:

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

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

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

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

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

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

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

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)[1]. The guess takes the following form: $guess^L = l^*_{it} - l^F_{it}$ and $guess^K = k^*_{it} - k^F_{it}$

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

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:

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

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

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

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

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

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, $a_t = A_1(L)\varepsilon_t$, $w_t = W(L)\varepsilon_t$, $r_t = R(L)\varepsilon_t$. 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.

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

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

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

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

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

where

$$\begin{gathered} l_t^{FA} = D(L)\varepsilon_t \\ k_t^{FA} = E(L)\varepsilon_t \\ l_{it}^{*A} = M^L(L)\varepsilon_t + N^L(L)\eta_{it}^L \\ k_{it}^{*A} = M^K(L)\varepsilon_t + N^K(L)\eta_{it}^K \end{gathered} \tag{F.1}$$

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

$$l_{it}^{FI} = k_t^{FI} = A_2(L)u_t$$
$$l_{it}^{*I} = k_{it}^{*I} = S(L)u_{it} + T(L)\psi_{it} \tag{F.2}$$

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:

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

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

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

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

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

subject to

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

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

where $\{m^K, m^L, n^K, n^L, s, t\}$ are the lag polynomial coefficients in equations (F.1) and (F.2).

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 :

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

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

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

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

Appendix G

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:

$$\max_{l_{it}, k_{it}} \left\{ e^{a_t} e^{a_{it}} k_{it}^{\alpha} l_{it}^{\delta} - w_t l_{it} - r_t k_{it} \right\} \tag{G.1}$$

The implied first order conditions are:

$$w_t = \delta e^{a_t} e^{a_{it}} k_{it}^{\alpha} l_{it}^{\delta - 1} \tag{G.2}$$

$$r_t = \alpha e^{a_t} e^{a_{it}} k_{it}^{\alpha - 1} l_{it}^{\delta} \tag{G.3}$$

Which implies :

$$\frac{w_t}{r_t} = \left(\frac{\delta}{\alpha}\right)\frac{k_{it}}{l_{it}} \tag{G.4}$$

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:

$$l_{it} = \left[\frac{\delta e^{a_t} e^{a_{it}} (\frac{\alpha w_t}{\delta r_t})^{\alpha}}{w_t}\right]^{\frac{1}{1-\alpha-\delta}} \tag{G.5}$$

$$k_{it} = l_{it}(\tfrac{w_t}{r_t}\tfrac{\alpha}{\delta}) \tag{G.6}$$

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:

$$C_t + K_{t+1} - (1-d)K_t = Y_t \tag{G.7}$$

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.

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

$$\psi \hat{L}_t + \gamma \hat{C}_t = \hat{w}_t \tag{G.8}$$

$$\hat{C}_t = E\left(\hat{C}_{t+1} - \frac{\hat{r}_{t+1}}{\gamma}\right) \tag{G.9}$$

$$\hat{w}_t = a_t + a_{it} + \alpha \hat{k}^F_{it} + (\delta - 1)\hat{l}^F_{it} \tag{G.10}$$

$$\hat{r}_t = a_t + a_{it} + (\alpha - 1)\hat{k}^F_{it} + \delta \hat{l}^F_{it} \tag{G.11}$$

$$\hat{k}^F_{it} - \hat{l}^F_{it} = \hat{w}_t - \hat{r}_t \tag{G.12}$$

$$\hat{y}^F_{it} = a_t + a_{it} + \alpha \hat{k}^F_{it} + \delta \hat{l}^F_{it} \tag{G.13}$$

$$\hat{l}^F_{it} = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - (1-\alpha)\hat{w}_t - \alpha \hat{r}_t) \tag{G.14}$$

$$\hat{k}^F_{it} = \frac{1}{1-\alpha-\delta}(a_t + a_{it} - \delta \hat{W}_t - (1-\delta)\hat{r}_t) \tag{G.15}$$

$$\hat{Y}_t = \frac{\bar{C}}{\bar{Y}}\hat{C}_t + \frac{\bar{K}}{\bar{Y}}\left(\hat{K}_{t+1} - (1-d)\hat{K}_t\right) \tag{G.16}$$

Aggregate Equilibrium Conditions under Perfect Information

By aggregating the firm-specific first order conditions I obtain 6 equations, three of which are equations $(G.8)$, (**??**) and (**??**), and 6 unknowns $\{L_t, K_{t+1}, C_t, Y_t, w_t, r_t\}$:

$$\hat{Y}_t = a_t + \alpha \hat{K}_t + \delta \hat{L}_t \tag{G.17}$$

$$\hat{L}_t = \frac{1}{1-\alpha-\delta}\left(a_t - (1-\alpha)\hat{w}_t - \alpha \hat{r}_t\right) \tag{G.18}$$

$$\hat{K}_t = \frac{1}{1-\alpha-\delta}(a_t - \delta \hat{W}_t - (1-\delta)\hat{r}_t) \tag{G.19}$$

Appendix H

Nash Bargaining Solution and Seller Heterogeneity

In this section, I solve a simple model with seller heterogeneity and derive the condition under which $m > m^*$ holds in equilibrium. I follow closely the methodology used in Lagos and Wright (2005).

Suppose agents differ in their productivity in the DM. That is, when an agent enters the DM as a seller, she can produce according to $c_\alpha(q) = \alpha q$. The productivity parameter α can take two values, $\{\alpha^H, \alpha^L\}$. The rest of the model is similar to Lagos and Wright.

In the CM, agents solve the following problem

$$W(m) = U(X) - H + \beta V(m')$$

s.t

$$X = H + \phi(m - m')$$

First order conditions and the Envelope condition are as follows

$$U'(X) = 1$$

$$\phi = \beta V'(m')$$

$$W'(m) = \phi$$

In the DM, with probability σ an agent can be a seller or a buyer, or doesn't trade at all. If the agent gets to be a buyer, she can be a a type α^H(low productivity) or a type α^L(high productivity) producer The value function of an agent with money holdings m entering the DM is

$$V_\alpha(m) = \sigma\left[-\alpha q(\tilde{m}, \alpha) + W(m + d(\tilde{m}, \alpha)\right] + \tag{H.1}$$

$$\sigma\left\{\varphi\left[u(q(m, \alpha^H) + W(m - d(m)\right] + (1-\varphi)\left[u(q(m, \alpha^L) + W(m - d(m)\right]\right\} + (1-2\sigma)W(m)$$

where φ denotes the share of type H agents. Equation (H.1) implies that the terms of trade in the DM depend only on the buyers money balances and the sellers productivity. Below, I solve the Nash Bargaining stage, where terms of trade are determined.

$$\max_{q, d \leq m} \left[u(q) - \phi d\right]^\theta \left[-\alpha q + \phi d\right]^{1-\theta}$$

The solution to this problem is

$$q(m,\alpha) = \begin{cases} q(m,\alpha) & \text{if } m < m^*(\alpha) \\ \\ q^*(\alpha) & \text{if } m \geq m^*(\alpha) \end{cases}$$

$$d(m,\alpha) = \begin{cases} m & \text{if } m < m^*(\alpha) \\ \\ m^* & \text{if } m \geq m^*(\alpha) \end{cases}$$

For cases when $m > m^*$ I have

$$u'(q^*) = \alpha \tag{H.2}$$

$$\phi m^* = (1-\theta)u(q^*) + \theta\alpha q^* \tag{H.3}$$

For cases when $m < m^*$,we have

$$\phi m = \frac{(1-\theta)u(q)\alpha + \alpha\theta u'(q)q}{\theta u'(q) + (1-\theta)\alpha} \tag{H.4}$$

$$d = m$$

where $q(m,\alpha)$ solves equation (H.4). At this point, similar to section 2.4.1, I need to prove that in equilibrium $m < m^*$ for both types and hence $d = m$. Focusing on this equilibrium, helps the tractability of the model since, otherwise there would be an additional source of heterogeneity. I proceed by showing the conditions under which $m < m^*$ holds for both types. Following Lagos and Wright, one can show that the value function (H.1) can be shown to be $V_t(m) = v_t(m) + \phi_t m + \max\{-\phi_t m' + \beta V_{t+1}(m')\}$

where

$$v_{\alpha,t}(m) \equiv \sigma \left\{ \varphi \left[u(q(m, \alpha^H) - \phi_t d_t(m) \right] + (1 - \varphi) \left[u(q(m, \alpha^L) - \phi_t d_t(m) \right] \right\}$$
$$+ \sigma \left\{ \phi_t d(\tilde{m}) - \tfrac{q}{\alpha} \right\} + U(X) - X$$

By repeated substitution, one can show that

$$V_{\alpha,t}(m_t) = v_{\alpha,t}(m_t) + \phi_t m_t + \sum_{j=t}^{\infty} \beta^{j-t} \max_{m_{j+1}} \left\{ -\phi_j m_{j+1} + \beta \left[v_{j+1}(m_{j+1}) + \phi_{j+1} m_{j+1} \right] \right\} \tag{H.5}$$

One needs to check under which conditions an equilibrium exists. In order to do this, I look at the slope of objective function (H.5) as $m_{t+1} \rightarrow m^*_{t+1}$ from below. In a model with sellers heterogeneity, there are two optimal quantities and money holdings, as shown in equations (H.2) and (2.35). Hence, I need check the equilibrium condition for each type. When a buyer comes across a high-productivity seller (α^L), the optimal amount of money held is less than if the buyer were to meet a low-productivity seller (α^H).

For a H-type (low productivity), the slope of (H.5) as $m_{t+1} \rightarrow m^*_{t+1}$ from below, is

$$\lim_{m_{t+1} \rightarrow m^*_{H,t+1}} \frac{\partial V_{H,t}(m_t)}{\partial m_{t+1}} = -\phi_t + \beta \phi_{t+1} + \beta \sigma \phi_{t+1} \Delta$$

where

$$\Delta = \varphi \frac{\alpha_H^2}{\theta(1-\theta) u_{qq}(q^*_H)(\alpha_H q^*_H - u) + \alpha_H^2} - 1$$

Except for $\phi_t = \beta \phi_{t+1}$ or $\theta = 1$, $\Delta < 0$, hence, in equilibrium, $m_{t+1} < m^*_{H,t+1}$.

For a L-type (high productivity), the slope of (H.5) as $m_{t+1} \to m^*_{L,t+1}$ from below, is

$$\lim_{m_{t+1}\to m^*_{L,t+1}} \frac{\partial V_{L,t}(m_t)}{\partial m_{t+1}} = -\phi_t + \beta\phi_{t+1} + \beta\sigma\phi_{t+1}\Psi$$

where

$$\begin{aligned}\Psi &= \varphi\frac{\alpha_L\left[\theta\alpha_L + (1-\theta)\alpha^H\right]}{\alpha_H(1-\theta)\theta u_{qq}(q^*_L)\left(\alpha_H q^*_L - u(q^*_L)\right) + \alpha_H\alpha_L\left[\theta\alpha_L + (1-\theta)\alpha^H\right]} \\ &\quad +(1-\varphi)\frac{\alpha_L^2}{\theta(1-\theta)\theta(1-\theta)u_{qq}(q^*_L)(\alpha_L q^*_L - u) + \alpha_L^2} - 1\end{aligned}$$

If $\Psi < 0$ then one can safely say that $m < m^*$ holds for both types of trade, between a buyer and high or low-type seller. Hence, in order to justify solving the model with sellers heterogeneity, under the situation when $m < m^*$, one has to restrict model's parameters such that $\Psi < 0$. We need

$$\varphi\frac{\alpha_L\left[\theta\alpha_L + (1-\theta)\alpha^H\right]}{\alpha_H(1-\theta)\theta u_{qq}(q^*_L)\left(\alpha_H q^*_L - u(q^*_L)\right) + \alpha_H\alpha_L\left[\theta\alpha_L + (1-\theta)\alpha^H\right]} +(1-\varphi)\frac{\alpha_L^2}{\theta(1-\theta)\theta(1-\theta)u_{qq}(q^*_L)(\alpha_L q^*_L - u) + \alpha_L^2} < 1$$

References

[1] Abras, Ana Luisa (2010). "Stable Firms and Unstable Wages: Theoory and Evidence on the Rise in Earnings Instability in the US Economy", working paper, University of Maryland

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

[3] 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.

[4] Aiyagari, S.Rao. 1994. "Uninsured idiosyncratic risk and Aggregate Saving", Quarterly Journal of economics Vol 109(3). Pp 659-684

[5] Albanesi, S (2002), "Inflation and inequality", working paper

[6] Algan, Y and Ragot, X (2006), "Monetary Policy with Heterogenous Agents and Credit Constraints", working paper

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

[8] Aruoba, B., C.Waller and R.Wright."Money and Capital: A Quantitative Analysis", manuscript, 2010.

[9] Bartelsman, E. J., J. C. Haltiwanger and S. Scarpetta, "Cross-Country Di¤erences in Productivity: The Role of Allocation and Selection," NBER Working Papers 15490, National Bureau of Economic Research, Inc, November 2009.

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

[11] Berentsen, .A, Camera.G and Ch, Waller (2005), "The distribution of Money Balances and the Nonneutrality of Money", International Economic Review, 46(2), 465-493.

[12] Berentsen, A, G. Camera and C.Waller (2006), "Money, Credit and Banking", forthcoming in Journal of Economic Theory.

[13] Bewley, Truman F. 1980. "The optimum Quantity of Money" in J.H Kareken and Wallace (eds), Models of Monetary Economies. Minneapolis: Federal Reserve Bank of Minneapolis

[14] Bhattacharya.J, J.Haslag and A. Martin (2005), "Heterogeneity, Redistribution and the Friedman Rule", International Economic Review, 46(2), 437-454.

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

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

[17] Cavalcanti, R. and N. Wallace (1999b). "A model of private bank-note issue"Review of Economic Dynamics, 2,104-136

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

[19] 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.

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

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

[22] 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.

[23] De Gregorio, J. (1996), "Borrowing Constraints, Human Capital Accumulation, and Growth," Journal of Monetary Economics, 37: 49–72.

[24] 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".

[25] Erosa, A and Ventura, G (2002) "On inflation as a Regressive Consumption Tax", Journal of Monetary Economics 49 p761-795

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

[27] 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

[28] Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. Review of Economic Studies 60, 35–52.

[29] 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.

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

[31] Huggett, Mark .1993. "The risk-free rate in heterogeneous-agent incomplete-insurance economies"Journal of Economic Dynamics and Control, 1993, vol. 17, issue 5-6, pages 953-969

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

[33] Imrohoroglu, A. 1992. "The welfare cost of inflation under imperfect insurance", Journal of Economic Dynamics and Control, 1992, vol. 16, issue 1, pages 79-91

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

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

[36] Kiyotaki, N, Wright, R. 1989 "On Money as a Medium of Exchange", Journal of Political Economy, Vol. 97, 1989, pp.927-954

[37] Kocherlakota, N. (2003). "Societal benefits of illiquid bonds", Journal of Economic Theory, 108, 179-193.

[38] Lagos, R. and R. Wright (2005) "A unified framework for monetary theory and policy analysis." Journal of Political Economy 113, 463—488.

[39] Lagos, Ricardo and Guillaume Rochetau (2008): "Money and Capital as Competing Media of Exchange," Journal of Economic Theory, 142, 247-258.

[40] Levine, (1991). "Asset trading mechanisms and expansionary monetary policy". Journal of Economic Theory, 54, 148-16

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

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

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

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

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

[46] 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.

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

[48] Molico, M and J.Chiu (2006), "Liquidity and the Welfare Cost of Inflation", manuscript

[49] Molico, M and Y.Zhang (2004),"Monetary Policy and the Distribution of Money and Capital", manuscript

[50] Molico, M. (2006) "The distribution of money and prices in search equilibrium." International Economic Review 47, 701-22

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

[52] Palacios-Huerta, I (2003), "An empirical Analysis of the Risk Properties of Human Capital Returns", American Economic Review 93(3), pp948-964.

[53] Pries, M (2001), "Uninsured Idiosyncratis Risk and Human Capital Accumulation", working paper

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

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

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

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

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

[59] 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

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

[61] Wallace N. "Whither Monetary Economics?" International Economic Review, 42: 847-869, November 2001.

[62] Wallace. N (2002), "General Features of monetary Models and their significance", prepared for "Swiss National Bank-Fed Cleveland Workshop on Monetary Economics"

[63] Zhu, Tao (2005), "An Overlapping generations model with Search", working paper