Essays in Dynamic Macroeconomics

Sang Seok Lee

Exeter College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy in Economics
at the University of Oxford in Trinity Term 2014

The approximate number of words: 44023

Declaration

I declare that this thesis represents my original work except where otherwise stated.

Dedication

To my parents, K. Y. Lee and M. S. Lim

To my siblings, S. H. Lee and K. E. Lee

Acknowledgements

First and foremost, I would like to thank Prof Martin Ellison who has supervised me throughout my graduate student life at Oxford. If it were not for his dedicated supervision, the completion of this thesis would not have been possible. I have benefited in countless ways from his intuition as well as his technical knowledge of economics.

In the process of writing this thesis, many people have provided me with helpful comments about various aspects of my work. I would like to express my gratitude to the following people: Guido Ascari, Paul Beaudry, Christopher Bowdler, Ian Crawford, Peter Eso, Irem Guceri, Refet Gurkaynak, Michal Horvath, Jooyoung Jeon, Thomas Lubik, Richard Mash, Antonio Mele, Kubilay Ozturk, Joe Pearlman, Andrew Pua, Amar Radia, Nicholas Woolley, and Francesco Zanetti. I also would like to thank Luca Guerrieri and Matteo Iacoviello for sharing their computer codes with me.

I am grateful to the Department of Economics at Oxford for providing me with financial assistance through the Doctoral Studentship Scheme. I am also grateful to Exeter College at Oxford for providing me with a wonderful working space.

Over the years, Lawrence Dacuycuy, Bombit Largoza, and Cesar Rufino have provided with me a seemingly unlimited supply of the Force. The young Padawan will always be in debt.

Writing this thesis would have been unbearable without support from my friends. Their company made it possible for me to focus on my research during turbulent periods in my life. In addition to those whose names were mentioned above, I would like to thank Silja Baller, Hrithik Bansal, Stephanie Chan, Marvin Chua, David Deregowski, Cosmina Dorobantu, Sarah Hegenbart, Seong Hoon Jeong, Fah Jirasavetakul, Chris Lakner, Banu Demir Pakel, Cavit Pakel, Maxim Polyakov, Maria Sveidahl, Steven Yit, and Anastasia Zolotareva for being there for me when I needed them. Kristel Chua deserves a special thank you.

Finally, I would like to thank my family for putting up with my absence in the past few years. I will make up for it in due time, I promise.

Essays in Dynamic Macroeconomics by Sang Seok Lee, Exeter College

A thesis submitted for the degree of Doctor of Philosophy in Economics at the University of Oxford in Trinity Term 2014

Abstract

This thesis is concerned with macroeconomic dynamics under various forms of uncertainty. Chapter 2 recognizes that the information flow from the interest rate is impeded when the nominal interest rate hits the zero lower bound. This impediment can (a) increase the duration of zero lower bound episodes and (b) bring about more persistent deflationary pressure. Moreover, it can make the exit from the zero lower bound disorderly.

Chapters 3 and 4 are concerned with dynamics of aggregate variables under Knightian Uncertainty. To overcome difficulties with expectations formation under Knightian Uncertainty, the agents follow Interactive Trial and Error Learning ($ITEL$) (Young, 2009; Pradelski and Young, 2012) to choose investment portfolios. This involves learning by occasionally experimenting with new actions even when the current action proves to be good. Two applications of $ITEL$ are presented. Chapter 3 deals with the growth of aggregate variables. The growth model can match several business cycle features of the US real aggregate wealth data. Chapter 4 considers a portfolio choice problem. The portfolio choice model can match the first two moments of the US real excess return of equity over bonds almost perfectly.

Chapter 5 explores "This Time Is Different Syndrome" of Reinhart and Rogoff (2009) in a setting where the agents are learning under Knightian Uncertainty. The agents are grouped into different generations and their models compete in terms of forecasting power. The predecessor's model is discarded together with the data set when its forecasting power is worse than the current generation's model. This loss of relevant data is rooted in focusing only on forecasting well in the short-run. By shifting the weight towards finding the true model of the economy, this problem can be substantially reduced.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is a collection of four essays about macroeconomic dynamics under various forms of uncertainty. The seminal work of Knight (1921) distinguishes risk from uncertainty as the following: whereas the former represents a situation in which the form of the unknown is well-known[1], the latter stands for a situation in which it is not. The latter has been known as Knightian Uncertainty in honor of his contribution. Up until recently, macroeconomists were primarily concerned with the concept of risk even though robust control macroeconomic theory, which started with a series of contributions by Lars Peter Hansen and Thomas Sargent[2], had been available as a modeling strategy to deal with Knightian Uncertainty. However, since the advanced economies entered deep recessions in 2008, mainstream macroeconomists have increasingly turned to the concept of Knightian Uncertainty to think about various macroeconomic issues.[3] Motivated by these recent developments, this thesis explores macroeconomic consequences of both risk and Knightian Uncertainty. Whereas the study of risk will be built upon Rational Expectations, the inquiry about Knightian Uncertainty will allow for departures from Rational Expectations.

Chapter 2 is concerned with risk. It develops on the idea that the zero lower bound

---

[1] The phrases such as "known unknown" and "unknown unknown" became popular when Donald Rumsfeld, then United States Secretary of Defense, used it to justify the US invasion of Iraq in 2003 when he was questioned that the motivation for the invasion was unsubstantiated.

[2] See Hansen and Sargent (2008) as well as the literature review in chapter 3 of this thesis.

[3] For instance, the Carnegie-Rochester Conference on Public Policy in 2014 was explicitly concerned with Knightian Uncertainty.

on the nominal interest rate is associated with a discrete jump in risk because the information flow from the interest rate is impeded at the zero lower bound. This mechanism is incorporated into the standard new Keynesian model which is the workhorse model in monetary economics. It will be shown that under some parametrizations, the impediment to the information flow at the zero lower bound (a) increases the duration of the zero lower bound periods and (b) generates more persistent deflationary pressure. Moreover, this problem makes the exit from the zero lower bound more disorderly: it is characterized by the oscillations of aggregate variables. This chapter also considers whether increased central bank transparency in the form of information revelation to public is beneficial.

In contrast to chapter 2, chapters 3, 4, and 5 are concerned with Knightian Uncertainty. In chapters 3 and 4, Knightian Uncertainty is defined as a situation in which the structure of the economic system is unknown to its agents. Because the agents do not know the structure, they encounter difficulty with forming expectations about the future which are necessary for deriving the optimal investment rules. To cope, they follow Interactive Trial and Error Learning ($ITEL$) (Young, 2009; Pradelski and Young, 2012) to choose their investment portfolios. $ITEL$ is a set of decision rules which involves learning by experimenting. Under $ITEL$, the agents occasionally experiment with new actions even when they are content with their current actions in order to improve their payoffs. Two macroeconomic applications of $ITEL$ are provided. Chapter 3 presents a growth model under Knightian Uncertainty. The model aims to match the business cycle features of the US real aggregate wealth data at quarterly frequency which are produced by the BBQ algorithm of Harding and Pagan (2002). This filtering method is particularly suitable for growth models because it uses both trend and cycle components of data. The growth model can match several business cycle features of the actual data such as the durations and the amplitudes of contractions and expansions. Chapter 4 considers a portfolio choice model under Knightian Uncertainty. The aim is to match the first two moments of the US real excess return of equity over bond at quarterly frequency which macroeconomists have been concerned with since the publication of Mehra and Prescott (1985). The portfolio choice model can match the first two moments of the actual data almost perfectly. Moreover, it can reproduce non-linearity in the actual data at a non-negligible frequency. The modeling strategy in these two chapters complements the standard approaches for dealing with model uncertainty such as robust control theory (see Hansen and Sargent, 2008) and econometric learning theory (see Evans and Honkapohja, 2001).

In chapter 5, Knightian Uncertainty takes the form of model specification uncertainty. Here, the agents are endowed with a set of models which contains the true model of the economy. This form of uncertainty is similar to the one used by robust control theory (see Hansen and Sargent, 2008). The aim of this chapter is to explore "This Time Is Different Syndrome" of Reinhart and Rogoff (2009) in a setting where learning takes place under Knightian Uncertainty. The model agents are grouped into different generations based on when they are born and their models compete over time in terms of forecasting power. The predecessor's model is discarded together with the data set used to estimate it when its forecasting power is worse than the current generation's model. Given that the true model of the economy is a stationary process, this leads to the loss of relevant and useful data. It will be shown that this is rooted in focusing only on forecasting well in the short-run. By shifting the weight towards finding the true model of the economy, this problem can be substantially reduced.

References

1. Evans, G. W. and Honkapohja, S. (2001). “Learning and expectations in macroeconomics,” Princeton Univeristy Press.
2. Hansen, L. P. and Sargent, T. J. (2008). “Robustness,” Princeton University Press.
3. Harding, D. and Pagan, A. (2002). “Dissecting the cycle: a methodological investigation,” Journal of Monetary Economics, Elsevier, vol. 49(2), pages 365-381, March.
4. Knight, F. H. (1921). “Risk, uncertainty, and profit,” Hart, Schaffner, and Marx.
5. Mehra, R. and Prescott, E. C. (1985). “The equity premium: A puzzle,” Journal of Monetary Economics, Elsevier, vol. 15(2), pages 145-161, March.
6. Pradelski, B. S. and Young, H. P. (2012). “Learning efficient Nash equilibria in distributed systems,” Games and Economic Behavior, Elsevier, vol. 75(2), pages 882-897, July.
7. Reinhart, C. M. and Rogoff, K. S. (2009). “This time is different: eight centuries of financial folly,” Princeton University Press.
8. Young, H. P. (2009). “Learning by trial and error,” Games and Economic Behavior, Elsevier, vol. 65(2), pages 626-643, March.

Chapter 2

Information Value of the Interest Rate and the Zero Lower Bound

Abstract

What are the consequences of information loss associated with the zero lower bound on the interest rate? This chapter recognizes that the interest rate contains valuable information about the state of the economy, and that this information flow is impeded when the nominal interest rate hits the zero lower bound. Under some parametrizations, this impediment to the information flow at the zero lower bound (a) increases the duration of zero lower bound episodes and (b) brings about more persistent deflationary pressure. Moreover, it makes the exit from the zero lower bound disorderly. These outcomes suggest that increased central bank transparency in the form of information revelation is beneficial. However, there exist parametrizations under which the increased central bank transparency is not beneficial.

2.1 Introduction

The recent experience by the US and the Eurozone has reignited an academic interest in the effect of the zero lower bound on the nominal interest rate on the aggregate economy. Following the seminal contribution of Jung, Teranishi, and Watanabe (2005), which was motivated by Japan's experience in the past decades, and Eggertsson and Woodford (2003), which was motivated by the US's experience in the early 2000s, many researchers have written about various issues regarding the effect of the zero lower bound lately. For instance, Adam and Billi (2007) and Nakov (2008) derive the optimal monetary policy under discretion and commitment subject to an occasionally-binding zero lower bound constraint and find that the commitment solution reduces welfare losses substantially relative to those under the discretionary solution. In relation to fiscal policy, Christiano, Eichenbaum, and Evans (2011) and Woodford (2011) show that the fiscal multiplier is larger than 1 at the zero lower bound. However, Aruoba and Schorfheide (2012) presents the result that the fiscal multiplier is not larger than 1 near the zero lower bound in the sunspot equilibrium. In their study of the new Keynesian model, Fernandez-Villaverde, Gordon, Guerron-Quintana, and Rubio-Ramirez (2012) find that imposing the zero lower bound alters the behavior of the model economy substantially. In a similar spirit, Nakata (2012) shows that increasing the variance of the discount factor shock reduces consumption, inflation, and output substantially more when the zero lower bound is binding than when it is not. In an open-economy setting, Bodenstein, Erceg, and Guerrieri (2009) demonstrate that the effects of foreign demand shocks on the home country are greatly amplified if the home country is constrained by the zero lower bound. In relation to exit strategy, Bianchi and Melosi (2014) find that the behavior of the economy at the zero lower bound depends on the beliefs about the exit strategy. Werning (2012) shows that the nominal interest rate should jump discretely upon exiting the zero lower bound under the optimal monetary policy. While this list is not exhaustive of developments in the literature, it is a good indicator of renewed interests in the zero lower bound problem by various researchers. This paper follows in their footsteps and considers (a) the relationship between the zero lower bound and the information flow from the nominal interest rate and (b) the effect of increasing the central bank transparency at the zero lower bound by modifying the information structure of the standard new Keynesian model.

The central mechanism in this paper is that the interest rate contains valuable information about the state of the economy and this information flow stops when the nominal interest rate hits the zero lower bound. As documented by Romer and Romer (2000) and Sims (2002), the Fed's Greenbook forecasts of inflation tend to be more accurate than the private sector's forecasts. Among other reasons, this is possibly because the Fed has access to a much larger information set than the private sector.[1] There is considerable empirical evidence in support of this notion. For instance, Kuttner (2001) shows that the bond rates' response to unanticipated changes in monetary policy is large and highly significant. Gurkaynak, Sack, and Swanson (2005) extends Kuttner (2001)'s analysis and demonstrates that the bond rates' and asset prices' response to the FOMC statements is also significant. Campbell, Evans, Fisher, and Justiano (2012) link these findings to forward guidance and interpret that the results above are due to the public's belief that the FOMC has private information about macroeconomic fundamentals. As for more direct evidence, Peek, Rosengren, and Tootell (1999) finds that the FOMC uses confidential bank supervisory information for guiding monetary policy. There is evidence that this applies to other central banks as well. For instance, Hubert (2013) shows that central bank inflation forecasts influence private sector inflation forecasts even when the former are not more accurate than the latter based on his study of five central banks (Canada, Japan, Sweden, Switzerland, and the UK). In a recent interview with BBC[2], Spencer Dale, the chief economist of the Bank of England, said the Bank did not possess any "special secrets" about the state of the economy. However, he also said

> "What we do - and we do it an awful lot - as well as look at the aggregate data published by our statistical office, is we spend an awful lot of time going up and down the country speaking to businesses and learning first-

---

[1] Romer and Romer (2000)'s conjecture is that this arises simply because the Fed devotes more resources to forecasting than the private sector.

[2] See http://www.bbc.co.uk/news/business-26167707.

> hand what's going on."

This type of informal surveying can be a source of information advantages for central banks given that many of them are better-equipped than their private sector counterparts for carrying out such activity. In this paper, the central bank has a larger information set than the private sector and uses an interest rate rule to set the nominal interest rate. As long as the nominal interest rate is outside the zero lower bound, the private sector can invert the interest rate rule and extract the missing piece of information which is informative about the state of the economy. However, this ceases to be the case at the zero lower bound because the interest rate rule is no longer invertible. This information problem at the zero lower bound complicates the signal extraction of the private sector and alters the dynamics of aggregate variables substantially through its effect on the expectation formation.

In this paper, the only piece of information that the private sector has to retrieve from the nominal interest rate is the current total factor productivity as only this is missing in its information set at the beginning of each time period. This choice is motivated by three considerations: (a) the total factor productivity is an important state variable for the private sector's expectation formation (through which the information problem at the zero lower bound on the nominal interest rate affects the dynamics of aggregate variables), (b) putting the total factor productivity at the center of the information problem is a promising way to think about the zero lower bound because the positive total factor productivity shocks, which are equivalent in their effect to the negative demand shocks for the new Keynesian models, are important for pushing the economy into it (as shown by Fernandez-Villaverde, Gordon, Guerron-Quintana, and Rubio-Ramirez (2012)), and (c) it is possible to incorporate the relationship between the information flow from the nominal interest rate and the zero lower bound into the standard new Keynesian model without major modifications when the information problem is associated with the total factor productivity.

The methodological novelty in this paper is the application of mathematical tools of censored-data microeconometrics to a dynamic macroeconomic model. Specifically, the expected value of the current total factor productivity when the zero lower bound on the nominal interest rate is binding is derived using the inverse Mills ratio. This expected value has an analytical expression which is highly tractable.

It will be shown that under some parametrizations, the information problem at the zero lower bound on the nominal interest rate makes (a) the duration of the zero lower bound periods longer and (b) the deflationary pressure in and out of the zero lower bound more persistent. Moreover, it makes the exit from the zero lower bound disorderly: it is accompanied by the fluctuations of the endogenous variables. Based on these observations, it will be established that increased central bank transparency in the form of announcing what the value of the current total factor productivity is can be beneficial because it resolves the information problem associated with the zero lower bound. This contributes to the literature that speaks in favor of increasing the central bank transparency such as Blinder (1998) and Woodford (2005). This type of information revelation can be considered as a form of forward guidance by which the central bank communicates the expected course of monetary policy to the private sector in order to manage the latter's expectations about the future. For a survey of other benefits of the increased central bank transparency, see Blinder et al. (2008). However, there exist parametrizations under which the increased central bank transparency is not beneficial. Even though the potential sub-optimality of increased central bank transparency has been reported by various researchers such as Cukierman (2009) and Gosselin, Lotz, and Wyplosz (2009), it is interesting that this can happen with the model in this paper as it does not have features that are especially conducive for making this point.

The paper is structured as follows: section 2.2 motivates the model in this paper; section 2.3 derives the model; section 2.4 discusses the solution method; section 2.5 establishes the main results; section 2.6 draws the policy implication and relates it to forward guidance; section 2.7 concludes.

2.2 Properties of the Standard New Keynesian Model under Different Parametrizations of the Interest Rate Rule

This section presents properties of the standard new Keynesian model under different parametrizations of the interest rate rule. The standard new Keynesian model here refers to a baseline new Keynesian model which can be found in standard monetary economics textbooks such as Walsh (2003), Woodford (2003), or Gali (2008). In this model, the private sector, which consists of households and firms, and the central bank have symmetric information sets (i.e., full information and identical) unlike in the following sections in which the latter has a larger information set than the former. The focus in this section is on volatilities of endogenous variables both in the absence and the presence of the zero lower bound on the nominal interest rate.

The standard new Keynesian model consists of equations[3]

$$x_t = E_t x_{t+1} - \frac{1}{\sigma}(\hat{R}_t - E_t \pi_{t+1}) + u_t \tag{2.1}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t \tag{2.2}$$

where E_t is a mathematical expectation based on the information set in t, x_t is an output gap in t, π_t is an inflation rate between $t-1$ and t, $\hat{R}_t$[4] is a nominal interest rate between t and $t+1$ which is set in t, u_t is a demand shock in t, σ is the coefficient of relative risk aversion, β is a discount factor, κ is a parameter which is a function of deep parameters. (2.1) is referred to as the IS Equation and (2.2) the new Keynesian Phillips curve in the literature.

The model above is usually closed by adding an equation that specifies how the nominal interest rate is set. Typically, researchers in the zero lower bound literature consider censored interest rate rules of the form $\hat{R}_t = \max[1 - \frac{1}{\beta}, \phi_x E_t x_{t-j} + \phi_\pi E_t \pi_{t-j}]$

---

[3] The full derivation can be found in Walsh (2003).

[4] The hat notation stands for the deviation of a variable from its steady state value.

where $1 - \frac{1}{\beta}$ is the nominal interest rate at the zero lower bound (as a deviation from its steady state value)[5] and $j \in \{-1, 0, 1\}$. The right element of this rule subsumes the backward-looking ($j = 1$), the current-looking ($j = 0$), and the forward-looking ($j = -1$) specifications which have been used by various researchers. Nakov (2008) shows that for given values of ϕ's, the backward-looking specification ($j = 1$) minimizes the welfare loss for the standard new Keynesian model with an occasionally-binding zero lower bound constraint on the nominal interest rate. Taking this as a starting point, this paper considers perturbations of the backward-looking specification above by $\phi_A \hat{A}_t$ or

$$\hat{R}_t = \max[1 - \frac{1}{\beta}, \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t] \tag{2.3}$$

as the interest rate rule in place where $\hat{A}_t$ is the total factor productivity in t. (2.3) is actually more general than being simple perturbations of the backward-looking specification above because the right element of (2.3) can be shown to be consistent with a more general specification of the form $\sum_{k=-1}^{1} \phi_{x,k} E_t x_{t-k} + \sum_{k=-1}^{1} \phi_{\pi,k} E_t \pi_{t-k}$ (refer to Appendix 2.8.1).[6] Then, the appearance of $\hat{A}_t$ in the right-hand side of (2.3) is a consequence of the presence of the current and forward-looking components in this more general specification.

It follows from the derivation of (2.1) that u_t takes the form

$$u_t = \frac{1+\eta}{\sigma+\eta} E_t \hat{A}_{t+1} - \frac{1+\eta}{\sigma+\eta} \hat{A}_t \tag{2.4}$$

where η is the inverse of the Frisch labor supply elasticity.[7] (2.4) makes clear that the total factor productivity is the source of the demand shock in this model. Finally, it

---

[5] Given that the steady state value of the (net) nominal interest rate R_t is $R^{ss} = \frac{1}{\beta} - 1$, $\hat{R}_t = R_t - R^{ss}$ implies that $\hat{R}_t = 1 - \frac{1}{\beta}$ when $R_t = 0$.

[6] This more general specification subsumes the backward-looking ($\phi_{x,-1} = \phi_{x,0} = \phi_{\pi,-1} = \phi_{\pi,0} = 0$), the current-looking ($\phi_{x,-1} = \phi_{x,1} = \phi_{\pi,-1} = \phi_{\pi,1} = 0$), and the forward-looking ($\phi_{x,0} = \phi_{x,1} = \phi_{\pi,0} = \phi_{\pi,1} = 0$) specifications as special cases.

[7] (2.4) emerges from rewriting the log-linearized equilibrium Euler equation $\hat{y}_t = E_t \hat{y}_{t+1} - \frac{1}{\sigma}(\hat{R}_{t+1} - E_t \pi_{t+1})$ in terms of the output gap $x_t = \hat{y}_t - \hat{y}_t^f$ by deducting $\hat{y}_t^f$ from its left-hand side and $y_t^f + (E_t y_{t+1}^f - E_t y_{t+1}^f) = y_t^f$ from its right-hand side in order to derive (2.1) where $\hat{y}_t$ is the sticky price equilibrium output and $\hat{y}_t^f = \frac{1+\eta}{\sigma+\eta} \hat{A}_t$ is the flexible price equilibrium output.

is assumed that

$$\hat{A}_t = \rho \hat{A}_{t-1} + \varepsilon_t \tag{2.5}$$

where $\varepsilon_t \sim IIDN(0, \sigma_\varepsilon)$ is a total factor productivity shock and ρ is an autoregressive coefficient.

In order to show how volatilites of the endogenous variables depend on $\phi_{\hat{A}}$ in (2.3), the standard new Keynesian model above is simulated 1000 times with each round being 2000 periods long (after discarding the initial 200 periods) for each value of $\phi_{\hat{A}}$ being considered here. The model is solved by Guerrieri and Iacoviello (2014)'s OccBin toolkit which can handle occasionally-binding constraints of various forms for linear rational expectations models (more on this toolkit in section 2.4). For simulation, $\sigma = 1, \eta = 1$, and $\beta = .99$ which are consistent with each time period being interpreted as a quarter.[8] $\kappa = .1717$ which follows from the parameter values above and the Calvo price stickiness parameter (Calvo, 1983) $\theta = .75$.[9] This value of θ implies that firms change their prices once a year on average which is backed up by some empirical evidence from micro data.[10] $\phi_x = .5$ and $\phi_\pi = 1.5$ which are commonly used values in the literature.[11] Finally, $\rho = .95$, and $\sigma_\varepsilon = .01$ which are usually considered to be consistent with each time period being a quarter.[12]

Table 2.1 gives the average standard deviations of the output gap, the inflation rate, and the nominal interest rate under different values of $\phi_{\hat{A}}$. It also provides the average frequency of the zero lower bound periods together with the maximum frequency. The reported results are averages of 1000 simulation rounds. Thus, the average frequency should be interpreted as the average across simulation rounds of the number of times the zero lower bound binds.[13] The table also provides the results for the case without the zero lower bound on the nominal interest rate.[14] The entries under "Without ZLB"

---

[8] For instance, see Gali (2008).

[9] $\kappa = \frac{(\sigma+\eta)(1-\theta)(1-\beta\theta)}{\theta}$ for the standard new Keynesian model. See Walsh (2003) for the derivation.

[10] For instance, see Alvarez et al. (2006) who survey the Euro area.

[11] For instance, see Nakov (2008).

[12] For instance, Cooley and Prescott (1995) use $\rho = .95$, and $\sigma_\varepsilon = .007$ for their quarterly calibration.

[13] Typically, several zero lower bound episodes of varying durations occur in each simulation round. The distribution of the durations of the zero lower bound episodes is provided in section 2.5.

[14] $\hat{R}_t = \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t$ in this case.

Table 2.1: Average Statistics under Different Values of $\phi_{\hat{A}}$

		$\phi_{\hat{A}} = .025$		$\phi_{\hat{A}} = .02$		$\phi_{\hat{A}} = .015$
SD (in %)	Without ZLB	With ZLB	Without ZLB	With ZLB	Without ZLB	With ZLB
σ_x	1.6874e-03	2.1764e-03	1.5749e-03	1.8668e-03	1.4625e-03	1.6300e-03
σ_π	2.8849e-03	3.0155e-03	2.6926e-03	2.7670e-03	2.5003e-03	2.5412e-03
$\sigma_{\hat{R}}$	4.2668e-03	4.24213-03	4.0784e-03	4.0607e-03	3.8905e-03	3.8784e-03
Ave Freq at ZLB		26.5348		19.7378		14.4040
Max Freq at ZLB		118		98		79
		$\phi_{\hat{A}} = .01$		$\phi_{\hat{A}} = .005$		$\phi_{\hat{A}} = 0$
SD (in %)	Without ZLB	With ZLB	Without ZLB	With ZLB	Without ZLB	With ZLB
σ_x	1.3500e-03	1.4402e-03	1.2375e-03	1.2830e-03	1.1250e-03	1.1462e-03
σ_π	2.3080e-03	2.3290e-03	2.1156e-03	2.1259e-03	1.9233e-03	1.9279e-03
$\sigma_{\hat{R}}$	3.7032e-03	3.6953e-03	3.5166e-03	3.5118e-03	3.3307e-03	3.3280e-03
Ave Freq at ZLB		10.4556		7.6540		5.6827
Max Freq at ZLB		59		48		34
		$\phi_{\hat{A}} = -.005$		$\phi_{\hat{A}} = -.01$		$\phi_{\hat{A}} = -.015$
SD (in %)	Without ZLB	With ZLB	Without ZLB	With ZLB	Without ZLB	With ZLB
σ_x	1.0125e-03	1.0215e-03	8.9997e-04	9.0331e-04	7.8747e-04	7.8856e-04
σ_π	1.7310e-03	1.7328e-03	1.5386e-03	1.5393e-03	1.3463e-03	1.3465e-03
$\sigma_{\hat{R}}$	3.1457e-03	3.1443e-03	2.9619e-03	2.9613e-03	2.7794e-03	2.7791e-03
Ave Freq at ZLB		4.3805		3.3309		2.7113
Max Freq at ZLB		27		19		15

σ_x, σ_π, and $\sigma_{\hat{R}}$ are the standard deviations of the output gap, the inflation rate, and the nominal interest rate. Ave Freq at ZLB is the average frequency of the zero lower bound periods based on simulation rounds in which the zero lower bound on the nominal interest rate binds at least for one period. Max Freq at ZLB is the maximum frequency of the zero lower bound periods in the set of simulation rounds in which the zero lower bound binds at least for one period. Without ZLB refers to the case in which the nominal interest rate is not constrained by the zero lower bound and With ZLB to the case in which it is constrained by the zero lower bound. The same sequences of shocks are used for simulation under different values of $\phi_{\hat{A}}$.

Table 2.2: Average Statistics for the Zero Lower Bound Periods under Different Values of $\phi_{\hat{A}}$

		$\phi_{\hat{A}} = .025$		$\phi_{\hat{A}} = .02$		$\phi_{\hat{A}} = .015$
SD (in %)	No ZLB	ZLB	No ZLB	ZLB	No ZLB	ZLB
σ_x	1.4401e-03	7.1467e-03	1.2640e-03	5.4076e-03	1.0815e-03	4.0284e-03
σ_π	1.2751e-03	3.6496e-03	1.0970e-03	2.7808e-03	9.2414e-04	2.0890e-03
$\sigma_{\hat{R}}$	1.5866-03	0	1.3806e-03	0	1.1711e-03	0
		$\phi_{\hat{A}} = .01$		$\phi_{\hat{A}} = .005$		$\phi_{\hat{A}} = 0$
SD (in %)	No ZLB	ZLB	No ZLB	ZLB	No ZLB	ZLB
σ_x	8.9906e-04	2.9424e-03	7.2987e-04	2.1889e-03	5.7683e-04	1.6162e-03
σ_π	7.4282e-04	1.5317e-03	5.9660e-04	1.1409e-03	4.6156e-04	8.3845e-04
$\sigma_{\hat{R}}$	9.7148e-04	0	8.1787e-04	0	6.8303e-04	0
		$\phi_{\hat{A}} = -.005$		$\phi_{\hat{A}} = -.01$		$\phi_{\hat{A}} = -.015$
SD (in %)	No ZLB	ZLB	No ZLB	ZLB	No ZLB	ZLB
σ_x	4.5870e-04	1.1701e-03	3.2874e-04	8.4942e-04	2.2489e-04	6.4393e-04
σ_π	3.5306e-04	6.0498e-04	2.5527e-03	4.2963e-04	1.8536e-04	3.1524e-04
$\sigma_{\hat{R}}$	5.5082e-04	0	3.9843e-03	0	3.1067e-04	0

The entries under ZLB are computed based on time periods and simulation rounds in which the zero lower bound on the nominal interest rate binds. The entries under No ZLB are calculated using the same time periods and simulation rounds as those under ZLB but without imposing the zero lower bound on the nominal interest rate. The results are obtained using the same simulation outputs as those for generating Table 2.1.

refer to this case whereas the entries under “With ZLB” refer to the case with the zero lower bound on the nominal interest rate as given by (2.3). The results clearly demonstrate the positive relationship between (a) $\phi_{\hat{A}}$ and the volatilities of the output gap and the inflation rate and (b) $\phi_{\hat{A}}$ and the frequency of the zero lower bound periods. (a) and (b) are related because the zero lower bound periods are characterized by jumps in volatilities of the output gap and the inflation rate. This is demonstrated in Table 2.2. It gives the average standard deviations of the output gap and the inflation rate conditional on the zero lower bound binding under “ZLB” and the counterfactuals that prevail in the absence of the zero lower bound under “No ZLB.” The entries under the latter are calculated using the same time periods and simulation rounds as those under the former. The results indicate that there are discrete jumps in the volatilities of the output gap and the inflation rate when the zero lower bound on the nominal interest rate binds.

What is an explanation for the positive relationship between $\phi_{\hat{A}}$ and the volatilities of the output gap and the inflation rate documented above? Intuitively, this is because the higher the value of $\phi_{\hat{A}}$, the higher the upward pressure on $\hat{R}_t$ (the nominal interest rate) when $\hat{A}_t$ (the total factor productivity) increases which is not desirable because an increase in $\hat{A}_t$ is associated with a decrease in x_t (the output gap) (as the demand shock is a decreasing function of $\hat{A}_t$) and a decrease in π_t (the inflation rate) (as the marginal cost of firms is a decreasing function of $\hat{A}_t$) for which a decrease rather than an increase in $\hat{R}_t$ is necessary for stabilization. Thus, the lower the value of $\phi_{\hat{A}}$ (i.e., moving in the direction of negative numbers), the more optimal the monetary policy is in the sense of bringing about lower volatilities of the output gap and the inflation rate. For instance, when the loss associated with monetary policy is measured by a quadratic function in the output gap and the inflation rate, a higher value of $\phi_{\hat{A}}$ implies a higher value of the loss.

In what follows, the standard new Keynesian model above will be modified. The modification centers around the idea that there is information asymmetry between the private sector and the central bank during the zero lower bound periods because the former cannot observe some information in the latter's information set during these periods. It will be shown that this modification has the effect of making the zero lower bound periods last longer while preserving the excess volatilities of the output gap and the inflation rate.

2.3 Model

2.3.1 Information Problem at the Zero Lower Bound

Let us start out by discussing the information problem associated with the zero lower bound on the nominal interest rate as this is the central feature of the model in this paper.

Unlike in the standard new Keynesian model of the previous section, now suppose that the central bank has a full information set but not the private sector[15] at the

---

[15] Refer to section 2.1 for the motivation of the idea that the central bank has a larger information

beginning of each time period. The central bank sets the nominal interest rate according to an interest rate rule which takes its information set as an input. If the interest rate rule is known to the private sector, the private sector can extract a useful signal about what is missing in its information set by inverting the central bank's interest rate rule and append it to its information set for the purpose of forming expectations. This is the sense in which the nominal interest rate movements have an additional informational value for the private sector.

Formally, let $\Omega_{1t}^{a} = (\omega_{1t}^{1}, \ldots, \omega_{n_1 t}^{1})$ be a vector of variables that appear in both the private sector and the central bank's information set at the beginning of t, $\Omega_{2t}^{a} = (\omega_{1t}^{2}, \ldots, \omega_{n_2 t}^{2})$ be a vector of variables that appear only in the central bank's information set at the beginning of t,[16] and $f : (\Omega_{1t}^{a}, \Omega_{2t}^{a}) \to \hat{R}_t$ be the central bank's linearized interest rate rule[17,18] whose functional form is $\sum_{i=1}^{n_1} \lambda_i^1 \omega_{it}^1 + \sum_{i=1}^{n_2} \lambda_i^2 \omega_{it}^2$. The superscript a stands for $ex - ante$ as the information sets above prevail before the private sector carries out its signal extraction exercise. Given that the nominal interest rate is set according to f, the private sector can extract the signal $\zeta_t = \sum_{i=1}^{n_2} \lambda_i^2 \omega_{it}^2$ uniquely by setting $\zeta_t = \hat{R}_t - \sum_{i=1}^{n_1} \lambda_i^1 \omega_{it}^1$ and use it to form its expectation.

However, the private sector cannot extract ζ_t uniquely when the nominal interest rate hits the zero lower bound because the nominal interest rate prescribed by f is not necessarily the nominal interest rate it observes: now, ζ_t is consistent with any point in the interval $(-\infty, 1 - \frac{1}{\beta} - \sum_{i=1}^{n_1} \lambda_i^1 \omega_{it}^1]$ where $1 - \frac{1}{\beta}$ is the value of the nominal interest rate at the zero lower bound (as a deviation from its steady state value). This information problem can have a significant consequence if ζ_t is important for the private sector's expectation formation.

In order to demonstrate the effect of the information problem at the zero lower

---

set than the private sector.

[16] The private sector may observe these variables with delays.

[17] This implies that the original interest rate rule is not necessarily restricted to be linear. However, working with a non-linear interest rate rule is beyond the scope of this paper. Given that solving a model with a non-linear solution method, which is what this paper does, is already more complicated than solving it with a linear solution method, introducing more non-linearity via non-linear equations makes things even more complicated.

[18] Because the equations used in this paper are products of linearization (hence linear themselves), it is acceptable to discuss the linearized interest rate rule straightaway without reference to the original interest rate rule.

bound, it is assumed in what follows that the current total factor productivity is the only variable that is missing in the private sector's information set at the beginning of each time period prior to the signal extraction exercise.[19] Assuming that only one variable is missing in the private sector's information set makes the interpretation of ζ_t above more straightforward as it is no longer a linear combination of multiple variables but a multiple of the missing variable from which the missing variable can be retrieved correctly (by dividing ζ_t by λ_1^2) given that ζ_t is uniquely extracted. Associating the information problem at the zero lower bound with the total factor productivity can produce dramatic effects on the endogenous variables because the total factor productivity is an important state variable for the private sector's expectation formation. Moreover, Fernandez-Villaverde, Gordon, Guerron-Quintana, and Rubio-Ramirez (2012)'s result, which says that the positive total factor productivity shocks (which are equivalent to negative demand shocks for the new Keynesian models) are important for pushing the economy into the interest rate zero lower bound for the new Keynesian models, suggests that putting the total factor productivity at the center of the information problem is a promising way to bring about a longer spell at the zero lower bound in this class of models. It will be shown that this is true under certain conditions. The use of the total factor productivity also allows incorporating the information problem without major modifications as it is already part of the standard new Keynesian model: there is no need to add another variable and equation.

Because only one variable is missing in the private sector's information set at the beginning of each time period, the information sets of the private sector and the central bank are "effectively symmetric" (i.e., full information and identical) when the zero lower bound on the nominal interest rate does not bind[20] but asymmetric when it binds.[21] What is meant by "effectively" will be clear once the information sets are

---

[19] One may wonder why the central bank has better knowledge of the total factor productivity than the private sector given that it is the firms that are more immediately affected by its variation. But, one should not confuse the firm-specific idiosyncratic productivity with the total factor productivity which is more closely associated with general macroeconomic conditions. The models in this paper do not feature the firm-specific idiosyncratic productivity.

[20] This is because the current total factor productivity can be retrieved correctly in this case as discussed in the previous paragraph.

[21] When there are more than one missing variables in the private sector's information set at the

discussed more explicitly in the following subsection.

2.3.2 Setup

2.3.2.1 Sequence of Movements

As suggested by the discussion in the previous subsection, the model agents move sequentially.[22] The central bank moves first and sets the nominal interest rate based on both Ω_1^a (the common information) and Ω_2^a (the central bank's additional information). After observing the nominal interest rate, the private sector moves and engages in the signal extraction exercise. Based on the outcome of this exercise, it updates its information set and carries out its actions which determine the output gap and the inflation rate.

2.3.2.2 Information Structure

The private sector observes the total factor productivity with at most one period lag. At the beginning of t, the intersection of the private sector and the central bank's information sets contains the output gap, the inflation rate, the total factor productivity, and the nominal interest rate up to $t-1$, i.e., $\Omega_{1t}^a = (x_{t-i}, \pi_{t-i}, \hat{A}_{t-i}, \hat{R}_{t-i})_{i=1}^{\infty}$. The variable that appears only in the central bank's information set at the beginning of t is the total factor productivity in t, i.e., $\Omega_{2t}^a = \hat{A}_t$. This information structure implies that $\hat{A}_t$ is known to the private sector at the beginning of $t+1$ whether it encounters the information problem in t or not.

As mentioned in the previous subsubsection, the central bank moves first and sets $\hat{R}_t$ using Ω_{1t}^a and Ω_{2t}^a. After observing $\hat{R}_t$, the private sector carries out the signal extraction exercise in order to figure out $\hat{A}_t$. Henceforth, let superscript p stand for

---

beginning of each time period, the information sets of the private sector and the central bank are asymmetric even when the zero lower bound on the nominal interest rate does not bind because what the private sector extracts uniquely from the signal extraction exercise is a linear combination of the missing variables, not the missing variables themselves as discussed in the previous paragraph. This linear combination is consistent with an infinite number of tuples of the missing variables.

[22] One can interpret the standard new Keynesian model in section 2.2 as being sequential because (2.3) implies that the central bank can move first and set the nominal interest rate without reference to the private sector's actions in that time period.

$ex - post$ which refers to time after the signal extraction exercise. When the zero lower bound on the nominal interest rate does not bind, the private sector can extract the underlying signal uniquely based on which $\hat{A}_t$ can be correctly retrieved. Then, $\Omega^p_{1t} = (x_{t-i}, \pi_{t-i}, \hat{A}_{t-i+1}, \hat{R}_{t-i+1})^{\infty}_{i=1}$ which appends $\hat{A}_t$ and $\hat{R}_t$ to Ω^a_{1t} and $\Omega^p_{2t} = \emptyset$ as the private sector and the central bank have the same information set in this case. This is what is meant by "effective symmetry" of the information sets in the previous subsection. When the zero lower bound binds, $\Omega^p_{1t} = (x_{t-i}, \pi_{t-i}, \hat{A}_{t-i}, \hat{R}_{t-i+1})^{\infty}_{i=1}$ which appends only $\hat{R}_t$ to Ω^a_{1t} and $\Omega^p_{2t} = \hat{A}_t$ which is identical to Ω^a_{2t}. This is because $\hat{A}_t$ is now consistent with a continuum of values. Thus, the information sets remain asymmetric even after the signal extraction exercise. In this case, the private sector has to work with $E^p_t \hat{A}_t = E[\hat{A}_t | \Omega^p_{1t}]$ instead of $\hat{A}_t$. The functional form of $E^p_t \hat{A}_t$ will be shown in subsubsection 2.3.2.5.

2.3.2.3 Central Bank

As discussed in section 2.2, the central bank sets the nominal interest rate using a censored interest rate rule[23] of the form

$$\hat{R}_t = \max[1 - \frac{1}{\beta}, \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t] \tag{2.6}$$

which is repeated here for convenience.[24] $1 - \frac{1}{\beta}$ is the value of the nominal interest rate at the zero lower bound (as a deviation from its steady state value).

2.3.2.4 Private Sector

Recall that when the nominal interest rate is above the zero lower bound, the private sector can retrieve $\hat{A}_t$ correctly by inverting the central bank's interest rate rule. Thus,

---

[23] Nakov (2008) shows that the interest rate zero lower bound binds too frequently when the nominal interest rate is set according to the optimal monetary policy under discretion or commitment. This is as high as 1/3 of the time which is too high relative to what is observed in the US and Japanese data. He says that "either that policy has not been conducted optimally (note that the frequency is much lower under the simple instrument rules) or that there may be other unmodeled costs associated with low or volatile interest rates, unrelated to the ability of the central bank to achieve its inflation and output-gap targets." The first argument motivates the use of an interest rate rule as a description of monetary policy in this paper.

[24] Refer to section 2.2 for the motivation behind this interest rate rule.

$E_t^p[\hat{A}_t] = E[\hat{A}_t|\Omega_{1t}^p] = \hat{A}_t$ as the $ex-post$ information set Ω_{1t}^p contains $\hat{A}_t$. In this case, (2.1) continues to be the relevant IS equation and (2.2) the relevant Phillips curve.[25] However, such is not true when the nominal interest rate hits the zero lower bound: $\hat{A}_t$ cannot be retrieved uniquely in this case. This subsubsection derives the IS equation and the Phillips curve which are consistent with this information problem at the zero lower bound.[26]

IS Equation

In the majority of new Keynesian models, the aggregate output is related to the total factor productivity and the aggregate hours that the households devote to the market employment (labor)[27] by

$$E_t^p \hat{y}_t = E_t^p \hat{A}_t + \hat{n}_t \tag{2.7}$$

where $\hat{y}_t$ is the aggregate output in t and $\hat{n}_t$ is the aggregate hours in t. When the zero lower bound on the nominal interest does not bind, $E_t^p \hat{y}_t = \hat{y}_t$ because $E_t^p \hat{A}_t = \hat{A}_t$. So, the firms know exactly how much output will be produced with their labor input in the period of production. However, this is not true when the zero lower bound binds because $E_t^p \hat{A}_t$ is not equivalent to $\hat{A}_t$. The functional form of $E_t^p \hat{A}_t$ for this case is derived later. Of course, the firms can figure out what $\hat{A}_t$ is after production because they find out what $\hat{y}_t$ is then. This is why the private sector observes the total factor productivity with at most one period lag (which was discussed in subsubsection 2.3.2.2).

The information problem also affects the aggregate consumption through the goods market equilibrium condition which takes the form

$$E_t^p \hat{c}_t = E_t^p \hat{y}_t \tag{2.8}$$

---

[25] E_t in (2.1) and (2.2) is identical to E_t^p in this case.

[26] The derivations here build on those which can be found in the standard macroeconomics textbooks such as Walsh (2003).

[27] There are models which add capital as a factor of production as well. For a recent survey of the literature, see Christiano, Trabandt, and Walentin (2010).

where $\hat{c}_t$ is the aggregate consumption in t. When the interest rate zero lower bound does not bind, $E_t^p \hat{c}_t = \hat{c}_t$ because $E_t^p \hat{y}_t = \hat{y}_t$. However, $E_t^p \hat{c}_t$ is not equivalent to $\hat{c}_t$ when the zero lower bound binds because $E_t^p \hat{y}_t$ is not equivalent to $\hat{y}_t$. In other words, the households do not know how much it will end up consuming because the firms do not know how much it will end up producing. (2.8) imposes the condition that the expected aggregate consumption should equal the expected aggregate output so that the households and the firms can decide how many hours of the households' employment is traded in the labor market.[28]

In line with (2.8), the log-linearized consumption Euler equation of the households becomes

$$E_t^p \hat{c}_t = E_t^p \hat{c}_{t+1} - \frac{1}{\sigma}(\hat{R}_t - E_t^p \pi_{t+1}). \tag{2.9}$$

Substituting (2.8) into (2.9) and applying the law of iterated expectations yields

$$E_t^p \hat{y}_t = E_t^p \hat{y}_{t+1} - \frac{1}{\sigma}(\hat{R}_t - E_t^p \pi_{t+1}). \tag{2.10}$$

In order to derive the IS equation similar to (2.1) using (2.10), it is necessary to define what the flexible price equilibrium output is. Here, it is defined as

$$E_t^p \hat{y}_t^f = \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_t \tag{2.11}$$

where $\hat{y}_t^f$ is the flexible price equilibrium output in t. (2.11) is an adaptation of the flexible price equilibrium output of the standard new Keynesian model to the setting here which involves replacing $\hat{A}_t$ with $E_t^p \hat{A}_t$ on the right-hand side and y_t^f with $E_t^p \hat{y}_t^f$ on the left-hand side. As above, $E_t^p \hat{y}_t^f = y_t^f$ when the zero lower bound on the nominal interest rate does not bind because $E_t^p \hat{A}_t = \hat{A}_t$ in this case. When the zero lower bound binds, $E_t^p \hat{y}_t^f$ is not equivalent to $\hat{y}_t^f$ because $E_t^p \hat{A}_t$ is not equivalent to $\hat{A}_t$. As usual, the relevant output gap $x_t = \hat{y}_t - \hat{y}_t^f$. Deducting $E_t^p y_t^f$ from the left-hand side of (2.10)

---

[28] This implicitly assumes that the information sets of the households and the firms are symmetric.

and $E_t^p y_t^f = E_t^p y_t^f + (E_t^p y_{t+1}^f - E_t^p y_{t+1}^f)$ from the right-hand side of (2.10) gives

$$E_t^p x_t = E_t^p x_{t+1} - \frac{1}{\sigma}(\hat{R}_t - E_t^p \pi_{t+1}) + u_t \tag{2.12}$$

where $u_t = E_t^p \hat{y}_{t+1}^f - E_t^p \hat{y}_t^f = \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_{t+1} - \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_t$ is the demand shock in t whose functional form follows from (2.11). (2.12) is the IS equation which is consistent with the information problem at the zero lower bound. When the zero lower bound on the nominal interest rate does not bind, $E_t^p x_t = x_t$ because $E_t^p \hat{y}_t = \hat{y}_t$ and $E_t^p \hat{y}_t^f = \hat{y}_t^f$ (since $E_t^p \hat{A}_t = \hat{A}_t$). In this case, (2.12) is equivalent to (2.1).[29] However, $E_t^p x_t$ is not equivalent to x_t when the zero lower bound binds because $E_t^p \hat{y}_t$ and $E_t^p \hat{y}_t^f$ are not equivalent to $\hat{y}_t$ and $\hat{y}_t^f$ respectively (since $E_t^p \hat{A}_t$ is not equivalent to $\hat{A}_t$). In such case, (2.12) is not equivalent to (2.1).

Recall that the private sector can figure out what $\hat{A}_t$ is after production when the zero lower bound on the nominal interest rate binds. This implies that x_t is known by the end of each time period as $\hat{y}_t$ and $\hat{y}_t^f$, which depend on $\hat{A}_t$, become known by the end of each time period and $x_t = \hat{y}_t - \hat{y}_t^f$. The equations above can be used to derive a relationship between x_t and $E_t^p x_t$ which is

$$\begin{aligned} x_t &= E_t^p x_t + (\hat{A}_t - E_t^p \hat{A}_t) - (\hat{y}_t^f - E_t^p \hat{y}_t^f) \\ &= E_t^p x_t + \frac{\sigma - 1}{\sigma + \eta}(\hat{A}_t - E_t^p \hat{A}_t). \end{aligned} \tag{2.13}$$

To derive (2.13), note that deducting (2.7) from $\hat{y}_t = \hat{A}_t + \hat{n}_t$ (post-production equivalent of (2.7)) and transposing $E_t^p y_t$ gives

$$\hat{y}_t = E_t^p \hat{y}_t + (\hat{A}_t - E_t^p \hat{A}_t). \tag{2.14}$$

Subtracting $\hat{y}_t^f$ from the left-hand side of (2.14) and $\hat{y}_t^f = \hat{y}_t^f + (E_t^p \hat{y}_{t+1}^f - E_t^p \hat{y}_{t+1}^f)$ from the right-hand side of (2.14) and using the definition of x_t yields the first line of (2.13).

---

[29] To be more specific, E_t in (2.1) and E_t^p in (2.12) are conditioned on the same information set. Then, $u_t = \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_{t+1} - \frac{1+\eta}{\sigma+\eta} \hat{A}_t$ which is equivalent to (2.4).

Finally, substituting (2.11) and $\hat{y}_t^f = \frac{1+\eta}{\sigma+\eta}\hat{A}_t$ (post-production equivalent of (2.11)) into the first line of (2.13) and collecting the like terms gives the second line of (2.13). (2.13) makes clear that the expectational error in the output gap $x_t - E_t^p x_t$ is driven by the expectational error in the total factor productivity $\hat{A}_t - E_t^p \hat{A}_t$. When the zero lower bound does not bind, (2.13) reduces to $E_t^p x_t = x_t$ as $E_t^p \hat{A}_t = \hat{A}_t$ in this case.[30]

Phillips Curve

In order to derive the Phillips curve similar to (2.2), note that new Keynesian models with the monopolistic competition (Dixit and Stiglitz, 1977) and the Calvo pricing (Calvo, 1983) by the intermediate good firms lead to

$$\pi_t = \beta E_t^p \pi_{t+1} + \tilde{\kappa} \hat{mc}_t \tag{2.15}$$

where $\hat{mc}_t$ is a real marginal cost of producing an intermediate good in t and $\tilde{\kappa}$ is a parameter that is a function of deep parameters. $\hat{mc}_t$ takes the form

$$\begin{aligned}
\hat{mc}_t &= \eta \hat{n}_t + \sigma E_t^p \hat{y}_t - E_t^p \hat{A}_t \\
&= \eta \hat{n}_t + (\eta E_t^p \hat{A}_t - \eta E_t^p \hat{A}_t) + \sigma E_t^p \hat{y}_t - E_t^p \hat{A}_t \\
&= (\eta + \sigma) E_t^p \hat{y}_t - (1+\eta) E_t^p \hat{A}_t \\
&= (\eta + \sigma)(E_t^p \hat{y}_t - \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_t) \\
&= (\eta + \sigma)(E_t^p \hat{y}_t - E_t^p \hat{y}_t^f) \\
&= (\eta + \sigma) E_t^p x_t. \qquad (2.16)
\end{aligned}$$

---

[30] When $\sigma = 1$, which is parametrization very commonly used in the literature, (2.13) implies that $E_t^p x_t = x_t$ irrespective of the expectational error in the total factor productivity $\hat{A}_t - E_t^p \hat{A}_t$, i.e., the output gap is self-fulfilling. However, this does not mean that the interest rate zero lower bound does not matter on x_t anymore (as the notational limitations may suggest misleadingly): because the form of E_t^p depends on whether the zero lower bound binds or not (as the structure of the $ex-post$ information set Ω_{1t}^p depends on it), x_t is still affected by the zero lower bound through its effect on $E_t^p x_t$.

The first line of (2.16) is a statement of the result that the real marginal cost is equal to the marginal rate of substitution between leisure and consumption in equilibrium. $E_t^p \hat{y}_t$ in the third line of (2.16) follows from (2.7) and $E_t^p \hat{y}_t^f$ in the fifth line of (2.16) from (2.11). Substituting (2.16) into (2.15) gives

$$\pi_t = \beta E_t^p \pi_{t+1} + \kappa E_t^p x_t \tag{2.17}$$

where $\kappa = (\eta + \sigma)\tilde{\kappa}$. (2.17) is the Phillips curve which is consistent with the information problem at the zero lower bound. As shown above, $E_t^p x_t = x_t$ when the zero lower bound on the nominal interest does not bind. In this case, (2.17) is equivalent to (2.2). However, this is not true when the zero lower bound binds as $E_t^p x_t$ is not equivalent to x_t in this case.

2.3.2.5 Signal Extraction

When the zero lower bound on the nominal interest rate does not bind, the private sector can retrieve the current total factor productivity $\hat{A}_t$ correctly based on the *ex−ante* information set $\Omega_{1t}^a = (x_{t-i}, \pi_{t-i}, \hat{A}_{t-i}, \hat{R}_{t-i})_{i=1}^{\infty}$, $\hat{R}_t$, and the backward-looking interest rate rule in (2.6) because they imply that

$$\hat{A}_t = \frac{1}{\phi_{\hat{A}}}\hat{R}_t - \frac{\phi_x}{\phi_{\hat{A}}}x_{t-1} - \frac{\phi_\pi}{\phi_{\hat{A}}}\pi_{t-1}. \tag{2.18}$$

But, this ceases to be the case when the zero lower bound binds because $\hat{A}_t$ is now consistent with a continuum of values. However, the assumption that the total factor productivity shock $\varepsilon_t \sim IIDN(0, \sigma_\varepsilon)$, in conjunction with (2.5) and (2.6), allows deriving the expression for $E_t^p \hat{A}_t$. In what follows, $\phi_{\hat{A}} > 0$ and $\phi_{\hat{A}} < 0$ are considered in turn.[31]

---

[31] When $\phi_{\hat{A}} = 0$, the signal extraction exercise is no longer necessary because the nominal interest rate no longer contains useful information about the current total factor productivity for the private sector.

When $\phi_{\hat{A}} > 0$

Let us start by noting that

$$\phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t \le 1 - \frac{1}{\beta}$$

$$\Rightarrow \varepsilon_t \le \frac{1 - \frac{1}{\beta}}{\phi_{\hat{A}}} - \frac{\phi_x}{\phi_{\hat{A}}} x_{t-1} - \frac{\phi_\pi}{\phi_{\hat{A}}} \pi_{t-1} - \rho \hat{A}_{t-1}. \tag{2.19}$$

The first line of (2.19) follows from (2.6) and the second line from substituting (2.5) into the first line and rearranging it. The right-hand side of the second line of (2.19) defines an upper bound on ε_t which is denoted by s_t henceforth. (2.5) and the *IIDN*ness of ε_t imply that $E_t^p \hat{A}_t$ takes the form

$$E_t^p \hat{A}_t = \rho \hat{A}_{t-1} + E[\varepsilon_t | \varepsilon_t \le s_t] \tag{2.20}$$

so the remaining task is figuring out what $E[\varepsilon_t|\varepsilon_t \le s_t]$ is. It is convenient to state that $E[\varepsilon_t|\varepsilon_t \le s_t] = E[\varepsilon_t|\frac{\varepsilon_t}{\sigma_\varepsilon} \le \frac{s_t}{\sigma_\varepsilon}] = \sigma_\varepsilon E[\frac{\varepsilon_t}{\sigma_\varepsilon}|\frac{\varepsilon_t}{\sigma_\varepsilon} \le \frac{s_t}{\sigma_\varepsilon}]$ because what follows builds on the fact that $\frac{\varepsilon_t}{\sigma_\varepsilon} \sim \Phi(\varepsilon)$ (standard normal distribution). Let $\tilde{\varepsilon}_t = \frac{\varepsilon_t}{\sigma_\varepsilon}$ and $\tilde{s}_t = \frac{s_t}{\sigma_\varepsilon}$ so that $E[\varepsilon_t|\varepsilon_t \le s_t] = \sigma_\varepsilon E[\tilde{\varepsilon}_t|\tilde{\varepsilon}_t \le \tilde{s}_t]$. It is a standard result that

$$E[\tilde{\varepsilon}_t|\tilde{\varepsilon}_t \le \tilde{s}_t] = \int_{-\infty}^{\tilde{s}_t} \tilde{\varepsilon}_t \frac{\phi(\tilde{\varepsilon}_t)}{\Phi(\tilde{s}_t)} d\tilde{\varepsilon}_t = \int_{-\infty}^{\tilde{s}_t} \frac{\frac{d}{d\tilde{\varepsilon}_t}(-\phi(\tilde{\varepsilon}_t))}{\Phi(\tilde{s}_t)} d\tilde{\varepsilon}_t = -\frac{\phi(\tilde{s}_t)}{\Phi(\tilde{s}_t)} \tag{2.21}$$

which implies that

$$E[\varepsilon_t|\varepsilon_t \le s_t] = -\sigma_\varepsilon \frac{\phi(\tilde{s}_t)}{\Phi(\tilde{s}_t)} \tag{2.22}$$

which is a non-linear function of the observables.[32] $\phi(.)$ and $\Phi(.)$ are the probability density function and the cumulative distribution function of the standard normal random variable. (2.22), which is referred to as the inverse Mills ratio in microeconometrics, serves as a central result in the analysis of censored data.[33] (2.22) is linearized

---

[32] This formula is applicable because $\tilde{s}_t$ is entirely predetermined which allows it to be treated as a constant in time period t.

[33] See Cameron and Trivedi (2005) for more details which include the derivation of (2.22).

in order to keep it consistent with the rest of the model. The most obvious point around which to linearize (2.22) is the value of $\tilde{s}_t$ for which the system enter the zero lower bound from the steady state: $\tilde{s}^{ss} = \frac{1}{\sigma_\varepsilon}(\frac{1-\frac{1}{\beta}}{\phi_{\hat{A}}})$. However, the linear approximation of (2.22) around $\tilde{s}^{ss}$ turns out to be a very poor description of (2.22).[34] Therefore, (2.22) is approximated around a constant $\tilde{s}^+$ which gives a more accurate result.[35] The linear approximation of (2.22) around $\tilde{s}^+$ is

$$-\sigma_\varepsilon \frac{\phi(\tilde{s}_t)}{\Phi(\tilde{s}_t)} \simeq \gamma_0^+ + \gamma_x^+ x_{t-1} + \gamma_\pi^+ \pi_{t-1} + \gamma_{\hat{A}}^+ \hat{A}_{t-1} \tag{2.23}$$

where

$$\begin{aligned}\gamma_0^+ &= -\sigma_\varepsilon \frac{\phi(\tilde{s}^+)}{\Phi(\tilde{s}^+)} - \sigma_\varepsilon \frac{\phi(\tilde{s}^+)(\tilde{s}^+\Phi(\tilde{s}^+) + \phi(\tilde{s}^+))}{\Phi(\tilde{s}^+)^2}(\tilde{s}^+ - \tilde{s}^{ss}),\\ \gamma_k^+ &= -\frac{\phi(\tilde{s}^+)(\tilde{s}^+\Phi(\tilde{s}^+) + \phi(\tilde{s}^+))}{\Phi(\tilde{s}^+)^2}\frac{\phi_k}{\phi_{\hat{A}}} \text{ for } k = x \text{and } \pi, \text{ and}\\ \gamma_{\hat{A}}^+ &= -\frac{\phi(\tilde{s}^+)(\tilde{s}^+\Phi(\tilde{s}^+) + \phi(\tilde{s}^+))}{\Phi(\tilde{s}^+)^2}\rho. \end{aligned} \tag{2.24}$$

Substituting (2.23) into (2.20) and collecting the like terms gives

$$E_t^p \hat{A}_t = \gamma_0^+ + \gamma_x^+ x_{t-1} + \gamma_\pi^+ \pi_{t-1} + (\rho + \gamma_{\hat{A}}^+)\hat{A}_{t-1}. \tag{2.25}$$

As mentioned above, the private sector substitutes $E_t^p \hat{A}_t$ in (2.25) for $\hat{A}_t$ when the zero lower bound on the nominal interest rate binds.

When $\phi_{\hat{A}} < 0$

Now,

$$\begin{aligned}\phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t &\le 1 - \frac{1}{\beta}\\ \Rightarrow \varepsilon_t &\ge \frac{1-\frac{1}{\beta}}{\phi_{\hat{A}}} - \frac{\phi_x}{\phi_{\hat{A}}} x_{t-1} - \frac{\phi_\pi}{\phi_{\hat{A}}}\pi_{t-1} - \rho\hat{A}_{t-1}.\end{aligned} \tag{2.26}$$

---

[34] $\tilde{s}^{ss}$ is the upper bound of the current total factor productivity shocks which make the economic system enter the zero lower bound from its steady state. However, such events are very unlikely: for the parametrizations used in this paper, $\Phi(\tilde{s}^{ss})$ is practically zero. In addition to this, the economic system enters the zero lower bound away from its steady state most of the time. Therefore, the linear approximation of (2.22) around $\tilde{s}^{ss}$is inappropriate for the purpose of simulation.

[35] Refer to subsections 2.5.1 and 2.5.2 which show that the approximation errors are small.

Following the same reasoning as above, with the right-hand-side of the second line of (2.26) being denoted by s_t again,

$$E[\varepsilon_t|\varepsilon_t \geq s_t] = \sigma_\varepsilon \frac{\phi(\tilde{s}_t)}{1-\Phi(\tilde{s}_t)} \tag{2.27}$$

where $\tilde{s}_t = \frac{s_t}{\sigma_\varepsilon}$. The linear approximation of (2.27) around a constant $\tilde{s}^-$ is

$$\sigma_\varepsilon \frac{\phi(\tilde{s}_t)}{1-\Phi(\tilde{s}_t)} \simeq \gamma_0^- + \gamma_x^- x_{t-1} + \gamma_\pi^- \pi_{t-1} + \gamma_{\hat{A}}^- \hat{A}_{t-1} \tag{2.28}$$

where

$$\begin{aligned}
\gamma_0^- &= \sigma_\varepsilon \frac{\phi(\tilde{s}^-)}{1-\Phi(\tilde{s}^-)} - \sigma_\varepsilon \frac{\phi(\tilde{s}^-)(-\tilde{s}^-(1-\Phi(\tilde{s}^-)) + \phi(\tilde{s}^-))}{(1-\Phi(\tilde{s}^-))^2}(\tilde{s}^- - \tilde{s}^{ss}), \\
\gamma_k^- &= -\frac{\phi(\tilde{s}^-)(-\tilde{s}^-(1-\Phi(\tilde{s}^-)) + \phi(\tilde{s}^-))}{(1-\Phi(\tilde{s}^-))^2} \frac{\phi_k}{\phi_{\hat{A}}} \text{ for } k = x \text{and } \pi \text{, and} \\
\gamma_{\hat{A}}^- &= -\frac{\phi(\tilde{s}^-)(-\tilde{s}^-(1-\Phi(\tilde{s}^-)) + \phi(\tilde{s}^-))}{(1-\Phi(\tilde{s}^-))^2}\rho.
\end{aligned} \tag{2.29}$$

Then,

$$E_t^p \hat{A}_t = \rho \hat{A}_{t-1} + E[\varepsilon_t|\varepsilon_t \geq s_t] = \gamma_0^- + \gamma_x^- x_{t-1} + \gamma_\pi^- \pi_{t-1} + (\rho + \gamma_{\hat{A}}^-)\hat{A}_{t-1} \tag{2.30}$$

which the private sector substitutes $\hat{A}_t$ for when the zero lower bound on the nominal interest rate binds.

2.3.3 Summary

Let us summarize the results in this section. The system of rational expectations equations which represents the economy is

1. $\hat{A}_t = \rho \hat{A}_{t-1} + \varepsilon_t$; $\varepsilon_t \sim IIDN(0, \sigma_\varepsilon)$
2. $\hat{R}_t = \max[1 - \frac{1}{\beta}, \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t]$

3. $$E_t^p \hat{A}_t = \begin{cases} \hat{A}_t & \text{if } \hat{R}_t > 1 - \frac{1}{\beta} \\ \gamma_0^+ + \gamma_x^+ x_{t-1} + \gamma_\pi^+ \pi_{t-1} + (\rho + \gamma_{\hat{A}}^+)\hat{A}_{t-1} & \text{if } \hat{R}_t \leq 1 - \frac{1}{\beta} \text{ and } \phi_{\hat{A}} > 0 \\ \gamma_0^- + \gamma_x^- x_{t-1} + \gamma_\pi^- \pi_{t-1} + (\rho + \gamma_{\hat{A}}^-)\hat{A}_{t-1} & \text{if } \hat{R}_t \leq 1 - \frac{1}{\beta} \text{ and } \phi_{\hat{A}} < 0 \end{cases}$$

4. $E_t^p x_t = E_t^p x_{t+1} - \frac{1}{\sigma}(\hat{R}_t - E_t^p \pi_{t+1}) + u_t$; $u_t = \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_{t+1} - \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_t$

5. $\pi_t = \beta E_t^p \pi_{t+1} + \kappa E_t^p x_t$

where $1 - \frac{1}{\beta}$ is the value of the nominal interest rate at the zero lower bound (as a deviation from its steady state value). The arrangement of the equations reflects the sequentiality in the model, except for the last two equations which are determined jointly by the private sector agents. The IS equation and the Phillips curve above are consistent with the information problem at the zero lower bound which is reflected in the use of $E_t^p x_t$ instead of x_t. As this notation may suggest, it is $E_t^p x_t$, not x_t, for which this system is solved for.

2.4 Solving the Model

The model in this paper is solved by modifying Guerrieri and Iacoviello (2014)'s OccBin toolkit which generates a non-linear solution to a system of rational expectations equations with occasionally-binding constraints.[36] Their solution algorithm builds on the solution techniques of Jung, Teranishi, and Watanabe (2005) and Eggertsson and Woodford (2003). It provides accurate solutions in the sense that they are very close to those obtained under dynamic programming. This is the case in spite of the limitation that their solution algorithm cannot capture precautionary motives associated with the possibility that constraints will be binding in the future as a result of unrealized future shocks. Its computational burden is also substantially lighter than one under higher quality methods which makes it attractive for stochastic simulations. In what follows, this algorithm is briefly described[37] in the context of the model here.

---

[36] There are alternative solution methods such as the cluster grid algorithm of Judd, Maliar, and Maliar (2011), the Smolyak collocation method of Fernandez-Villaverde, Gordon, Guerron-Quintana, and Rubio-Ramirez (2012), and the shadow price shocks approach of Holden and Paetz (2012).

[37] See Bodenstein, Erceg, and Guerrieri (2009) for the details of the algorithm.

In its simplest form, Guerrieri and Iacoviello (2014)'s algorithm generates the impulse response functions by (a) conjecturing the last period in which the zero lower bound on the nominal interest rate binds, (b) solving the model backward from the last period in which the zero lower bound binds to the initial period in which the zero lower bound binds to generate a time-dependent solution, and (c) validating whether this solution is consistent with the conjecture about the last period in which the zero lower bound binds. These steps are repeated until convergence. Their algorithm can also handle more complicated dynamics such as an oscillation in and out of the zero lower bound for the in-between periods. As mentioned above, the algorithm can be used for stochastic simulations as well.

In this paper, Guerrieri and Iacoviello (2014)'s algorithm is modified so that it is consistent with the information structures of the model, i.e., it should be that the correct values of $\hat{A}_t$ (total factor productivity) and x_t (output gap) are not known before production but known after production when the zero lower bound on the nominal interest rate binds in t. It is shown in subsections 2.5.1 and 2.5.2 that the approximation errors are small in spite of this modification.

2.5 Results

In addition to the model in section 2.3 which features the information problem at the zero lower bound on the nominal interest rate ("With ZLB & IP"), the standard new Keynesian model with and without the zero lower bound on the nominal interest rate ("With ZLB" and "Without ZLB" respectively), whose properties were discussed in section 2.2 under the same names, are used for stochastic simulations. Recall that whereas the private sector and the central bank have the same (full) information set at all times under "Without ZLB" and "With ZLB," the central bank has a larger information set than the private sector under "With ZLB & IP" which is at the root of the information problem at the zero lower bound. These models are subject to a two standard deviation shock to the total factor productivity initially ($\varepsilon_1 = .02$)[38] in the neighborhood

---

[38] The results in this section hold under smaller shocks as well.

of the zero lower bound which is obtained from the actual simulation of "With ZLB" under a randomly-generated sequence of the total factor productivity shocks.[39] The motivation behind starting in the neighborhood of the zero lower bound rather than from the steady state is that the zero lower bound periods are a phenomenon that occurs sufficiently away from the steady state. Because the models under different values of $\phi_{\hat{A}}$ are different models, they enter the zero lower bound from different conditions under the same sequence of the total factor productivity shocks. For each value of $\phi_{\hat{A}}$, "Without ZLB", "With ZLB," and "With ZLB & IP" are simulated 1000 times each (by applying randomly-generated sequences of the total factor productivity shocks to the models above after the initial period) in order to illustrate the distribution of possible paths of the output gap, the inflation rate, and the nominal interest rate. From this, results which hold across different values of $\phi_{\hat{A}}$ will emerge. The same sequences of the total factor productivity shocks are used across different models under different values of $\phi_{\hat{A}}$. The results reported in this section are robust to using different neighborhoods of the zero lower bound which are obtained from using different sequences of the total factor productivity shocks.[40]

For simulation, $\sigma = 1$, $\eta = 1$, $\beta = .99$, $\kappa = .1717$, $\phi_x = .5$, $\phi_\pi = 1.5$, $\rho = .95$, and $\sigma_\varepsilon = .01$ which are identical to those under section 2.2.[41] The results in this section are qualitatively robust to using different values for these parameters. $\tilde{s}^+ = .02$ and $\tilde{s}^- = -.02$ as these values turn out to be good for the numerical stability. It will be shown that these values keep the private sector's expectational errors about the current total factor productivity at the zero lower bound small as well which prevents the results in this section being driven by large expectational mistakes. $\phi_{\hat{A}}$ ranges from -.015 to .025 with .005 increments.[42]

As shown in subsubsection 2.3.2.5, the mathematical structure of the new Keyne-

---

[39] This way of obtaining the neighborhood of the zero lower bound is acceptable because the three models are practically identical outside the zero lower bound.

[40] This is because the new Keynesian models enter the zero lower bound from similar situations in which the output gap and the inflation rate are sufficiently negative.

[41] Refer to section 2.2 for the explanations.

[42] For values above .025 (.027 onward), it is difficult to achieve the numerical stability uniformly. For values below -.015, the zero lower bound periods occur extremely infrequently.

sian model with the information problem at the zero lower bound ("With ZLB & IP") changes when the sign of $\phi_{\hat{A}}$ changes. For the sake of conservatism, this section will report results under the positive and the negative $\phi_{\hat{A}}$'s separately.[43]

Tables 2.3 and 2.4 report the median paths[44] of the output gap (x), the inflation rate (π), and the nominal interest rate (in level) (R) under negative and positive values of $\phi_{\hat{A}}$ respectively when the three models are subject to the two standard deviation shock to the total factor productivity initially. The numbers correspond to time periods with 1 being the first period inside the zero lower bound.

Let us start by discussing the entries under "Without ZLB." As expected, the nominal interest rate falls below zero initially without the zero lower bound on it. The output gap, the inflation rate, and the nominal interest rate smoothly return to their steady state values over time after the initial shock.[45] This result holds whether $\phi_{\hat{A}}$ is positive or negative.

Things are not quite as smooth with the zero lower bound in place. Comparing "With ZLB & IP" with "With ZLB", two interesting results emerge. First, the duration of the zero lower bound periods, which is provided under Dur, is longer with the information problem at the zero lower bound for the positive $\phi_{\hat{A}}$'s. However, this is not uniformly true for the negative $\phi_{\hat{A}}$'s: once $\phi_{\hat{A}}$ becomes sufficiently negative, the result gets overturned. The tables suggest that on the both sides of zero for $\phi_{\hat{A}}$, there is a negative relationship between $\phi_{\hat{A}}$ and the extra periods inside the zero lower bound due to the information problem. Second, while the inflation rate reacts more initially in the absence of the information problem, the deflationary pressure is more persistent in its presence as evidenced by the trajectory of the inflation rate inside and immediately

---

[43] For a given sequence of the total factor productivity shocks, even though the models whose $\phi_{\hat{A}}$'s have the same sign tend to enter the zero lower bound at the same point of the sequence, the point of the entry differs under different signs of $\phi_{\hat{A}}$. This makes applying the same neighborhood of the zero lower bound (in terms of its association with a particular point in a sequence of the total factor productivity shocks) under different signs of $\phi_{\hat{A}}$ difficult. Thus, it is not fair to draw conclusions by pooling the results under different signs of $\phi_{\hat{A}}$ together from the beginning (even though some conclusions will be drawn irrespective of the sign of $\phi_{\hat{A}}$).

[44] By median path, what is meant is the evolution of the median of a variable's distribution at each point in time.

[45] x in time period 10 is smaller than the one in time period 1 for $\phi_{\hat{A}} = .025$ because of the U-shaped dynamics of the output gap. This occurs for lower values of $\phi_{\hat{A}}$ as well but it is not captured by the tables because it is briefer than 10 periods for those values.

Table 2.3: Median Paths of Endogenous Variables under Negative Values of $\phi_{\hat{A}}$

$\phi_{\hat{A}} = -.005$					
Without ZLB	1	10	20	30	40
x (%)	-.14	-.16	-.09	-.05	-.03
π (%)	-.69	-.45	-.27	-.15	-.08
R (%)	-.46	.17	.50	.72	.86
With ZLB	1	10	20	30	40
x (%)	-1.53	-.14	-.09	-.05	-.03
π (%)	-1.16	-.43	-.27	-.15	-.08
R (%)	0	.13	.50	.72	.86
With ZLB & IP	1	10	20	30	40
x (%)	-.38	-.14	-.09	-.05	-.03
π (%)	-.56	-.48	-.27	-.15	-.08
R (%)	0	0	.50	.72	.86

$\phi_{\hat{A}} = -.01$					
Without ZLB	1	10	20	30	40
x (%)	-.22	-.14	-.08	-.04	-.02
π (%)	-.63	-.40	-.24	-.13	-.07
R (%)	-.23	.22	.53	.74	.87
With ZLB	1	10	20	30	40
x (%)	-.91	-.13	-.08	-.04	-.02
π (%)	-.84	-.39	-.24	-.13	-.07
R (%)	0	.21	.53	.74	.87
With ZLB & IP	1	10	20	30	40
x (%)	-.29	-.12	-.07	-.04	-.02
π (%)	-.54	-.40	-.24	-.13	-.07
R (%)	0	.18	.53	.74	.87

$\phi_{\hat{A}} = -.015$					
Without ZLB	1	10	20	30	40
x (%)	-.22	-.12	-.07	-.04	-.02
π (%)	-.55	-.35	-.21	-.12	-.06
R (%)	-.50	-.12	-.07	-.04	-.02
With ZLB	1	10	20	30	40
x (%)	-.50	-.12	-.07	-.04	-.02
π (%)	-.63	-.35	-.21	-.12	-.06
R (%)	0	.26	.56	.76	.88
With ZLB & IP	1	10	20	30	40
x (%)	-.20	-.12	-.07	-.04	-.02
π (%)	-.50	-.35	-.21	-.12	-.06
R (%)	0	.26	.56	.76	.88

$\phi_{\hat{A}} = -.005$	Dur
With ZLB	8
With ZLB & IP	10
$\phi_{\hat{A}} = -.01$	Dur
With ZLB	7
With ZLB & IP	7
$\phi_{\hat{A}} = -.015$	Dur
With ZLB	5
With ZLB & IP	4

Without ZLB corresponds to the standard new Keynesian model without the zero lower bound on the nominal interest rate, With ZLB to the standard new Keynesian model with the zero lower bound on the nominal interest rate, and With ZLB & IP to the new Keynesian model with the information problem at the zero lower bound on the nominal interest rate (which was derived in section 2.3). x stands for the output gap, π for the inflation rate, and R for the nominal interest rate in level (not as a deviation from its steady state value). 1, 10, 20, 30, and 40 are time periods with 1 being the period in which the economy enters the zero lower bound. The entries below the numbers are the values of the endogenous variables in the corresponding time periods. Dur denotes the duration of the zero lower bound periods. The reported figures are the medians of 1000 simulation rounds.

Table 2.4: Median Paths of Endogenous Variables under Positive Values of $\phi_{\hat{A}}$

$\phi_{\hat{A}} = .025$						$\phi_{\hat{A}} = .02$					
Without ZLB	1	10	20	30	40	Without ZLB	1	10	20	30	40
x (%)	-.16	-.20	-.12	-.06	-.03	x (%)	-.22	-.19	-.11	-.05	-.03
π (%)	-.91	-.59	-.36	-.19	-.10	π (%)	-.86	-.55	-.34	-.18	-.10
R (%)	-.62	.12	.46	.71	.86	R (%)	-.43	.16	.48	.72	.86
With ZLB	1	10	20	30	40	With ZLB	1	10	20	30	40
x (%)	-2.49	-.16	-.11	-.05	-.03	x (%)	-1.74	-.16	-.10	-.05	-.03
π (%)	-1.77	-.53	-.35	-.19	-.10	π (%)	-1.39	-.51	-.33	-.18	-.09
R (%)	0	0	.46	.71	.86	R (%)	0	0	.48	.72	.86
With ZLB & IP	1	10	20	30	40	With ZLB & IP	1	10	20	30	40
x (%)	-.07	-.08	-.23	.10	-.07	x (%)	-.07	-.26	-.13	-.07	-.03
π (%)	-.94	-.82	-.69	-.24	-.12	π (%)	-.86	-.72	-.37	-.18	-.10
R (%)	0	0	0	.10	.78	R (%)	0	0	.36	.70	.86
$\phi_{\hat{A}} = .015$						$\phi_{\hat{A}} = .01$					
Without ZLB	1	10	20	30	40	Without ZLB	1	10	20	30	40
x (%)	-.25	-.18	-.10	-.05	-.03	x (%)	-.25	-.16	-.10	-.05	-.03
π (%)	-.80	-.51	-.31	-.17	-.09	π (%)	-.74	-.48	-.29	-.16	-.08
R (%)	-.31	.20	.51	.73	.87	R (%)	-.21	.24	.53	.74	.88
With ZLB	1	10	20	30	40	With ZLB	1	10	20	30	40
x (%)	-1.18	-.15	-.09	-.05	-.03	x (%)	-.78	-.15	-.09	-.04	-.02
π (%)	-1.11	-.49	-.31	-.17	-.09	π (%)	-.90	-.45	-.29	-.15	-.08
R (%)	0	.14	.51	.73	.87	R (%)	0	.20	.53	.74	.88
With ZLB & IP	1	10	20	30	40	With ZLB & IP	1	10	20	30	40
x (%)	-.07	-.07	-.10	-.05	-.03	x (%)	-.08	-.11	-.09	-.05	-.03
π (%)	-.80	-.51	-.32	-.17	-.09	π (%)	-.75	-.50	-.29	-.15	-.08
R (%)	0	0	.47	.73	.87	R (%)	0	0	.53	.74	.88
$\phi_{\hat{A}} = .005$						$\phi_{\hat{A}} = .025$					Dur
Without ZLB	1	10	20	30	40	With ZLB					11
x (%)	-.26	-.15	-.09	-.04	-.02	With ZLB & IP					28
π (%)	-.69	-.44	-.27	-.14	-.07	$\phi_{\hat{A}} = .02$					Dur
R (%)	-.11	.28	.56	.76	.89	With ZLB					10
With ZLB	1	10	20	30	40	With ZLB & IP					16
x (%)	-.51	-.14	-.08	-.04	-.02	$\phi_{\hat{A}} = .015$					Dur
π (%)	-.75	-.42	-.26	-.14	-.07	With ZLB					8
R (%)	0	.26	.56	.76	.89	With ZLB & IP					11
With ZLB & IP	1	10	20	30	40	$\phi_{\hat{A}} = .01$					Dur
x (%)	-.08	-.14	-.08	-.04	-.02	With ZLB					6
π (%)	-.70	-.44	-.27	-.14	-.07	With ZLB & IP					9
R (%)	0	.22	.56	.76	.89	$\phi_{\hat{A}} = .005$					Dur
						With ZLB					5
						With ZLB & IP					6

Refer to Table 2.3 for the explanations.

outside the zero lower bound. This result holds irrespective of the sign of $\phi_{\hat{A}}$.[46]

In order to achieve a better understanding of the observations above, in what follows the three models will be analyzed more thoroughly. Two values of $\phi_{\hat{A}}$ are considered in turn: .025 and -.01. This choice allows demonstrating how the sign of $\phi_{\hat{A}}$ affects the dynamics of the models. Appendix 2.8.2 provides the results under other values of $\phi_{\hat{A}}$ in order to complement the analysis below.

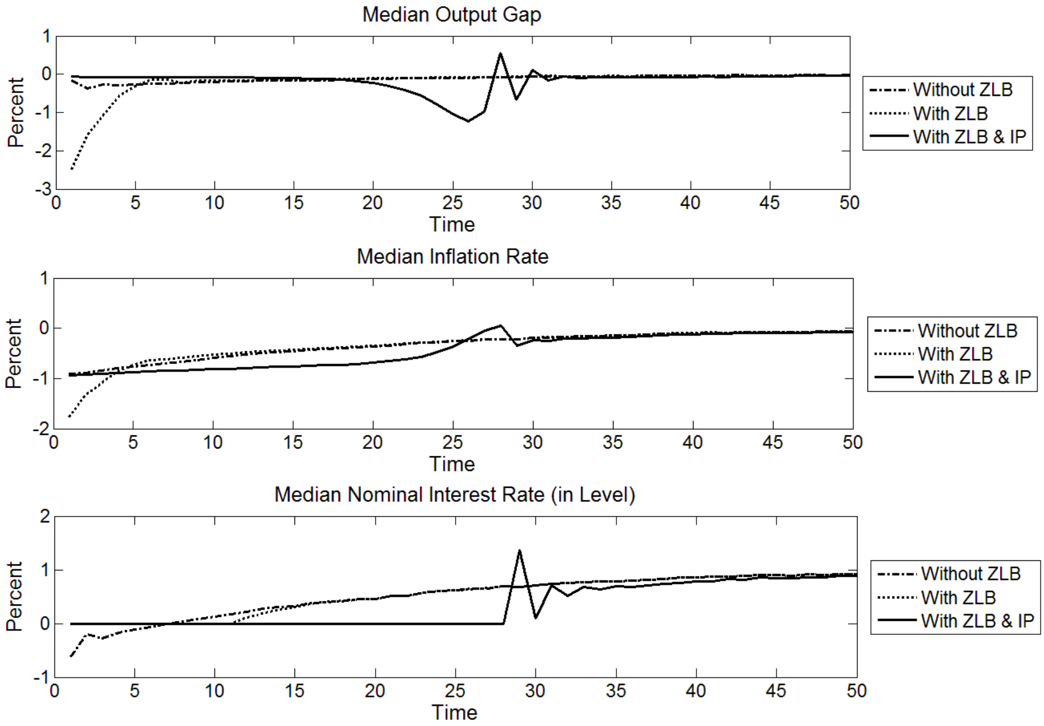
2.5.1 $\phi_{\hat{A}} = .025$

Figure 2.1 gives full plots of the median paths of the output gap, the inflation rate, and the nominal interest rate (in level) which were partially provided in Table 2.4 above. Recall that these are generated by subjecting the standard new Keynesian model without the zero lower bound on the nominal interest rate ("Without ZLB"), the standard new Keynesian model with the zero lower bound ("With ZLB"), and the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP") to the two standard deviation shock in the neighborhood of the zero lower bound. The figure confirms the observations above that (a) the duration of the zero lower bound periods is longer and (b) the deflationary pressure is more persistent with the information problem at the zero lower bound. Moreover, the figure suggests that the exit from the zero lower bound is disorderly: it is accompanied by the oscillations of the endogenous variables. The figure also shows that the information problem alters the dynamics of the output gap and the inflation rate during the zero lower bound periods substantially.

---

[46] Such is the case for $\phi_{\hat{A}} = -.015$ as well but this is not shown on Table 2.3 because this lasts for less than 10 periods.

Figure 2.1: Median Paths of Endogenous Variables



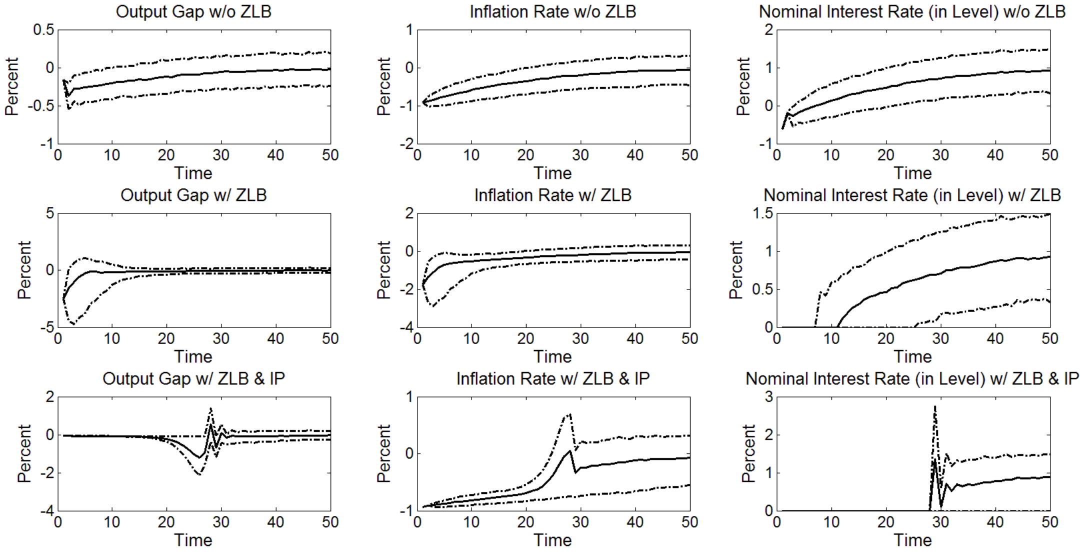
The figure provides the median paths of the output gap, the inflation rate, and the nominal interest rate (in level) based on 1000 simulation rounds. Without ZLB corresponds to the standard new Keynesian model without the zero lower bound on the nominal interest rate, With ZLB to the standard new Keynesian model with the zero lower bound, and With ZLB & IP to the new Keynesian model with the information problem at the zero lower bound. $\phi_{\hat{A}} = .025$.

Figure 2.2 provides the 10th percentiles and the 90th percentiles (broken lines) of the output gap, the inflation rate, and the nominal interest rate (in level) in addition to their medians (unbroken lines). It shows that for the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP"), a lot of actions come from the dynamics of the inflation rate. This is supported by the result that the inflation rate remains strongly and significantly away from zero even after many time periods (as the 90% confidence interval does not contain zero). Intuitively, this is because (a) the interest rate rule in (2.6) responds more aggressively to the inflation rate and (b) this is reflected in the coefficient of the inflation rate being much larger than the coefficients of other variables in the conditional expectation of the current total factor productivity at the zero lower bound in (2.25)[47]. This implies that the

---

[47] For the parameter values used in this paper, $\gamma_0^+ = -.0487$, $\gamma_x^+ = -2.2715$, $\gamma_\pi^+ = -6.8144$, and

Figure 2.2: Distributions of Endogenous Variables



The unbroken lines are for the medians of the variables and the broken lines are for the 10th and the 90th percentiles of the variables. The labeling conventions are identical to Figure 2.1. $\phi_{\hat{A}} = .025$.

private sector's expectations about the future at any point in time during the zero lower bound periods reacts more sensitively to the inflation rate and this feeds back into the inflation rate. Thus, (2.25) can be thought of as an "expectations accelerator."

As shown by Koop, Pesaran, and Potter (1996), analyzing non-linear models with the ordinary impulse functions can be misleading as their suitability is limited to some linear models. Because the models in this paper ("With ZLB" and "With ZLB & IP") are non-linear, the same caution should apply here. Therefore, the analysis above is complemented by the generalized impulse response functions which are more suitable than the ordinary impulse response functions. Following Koop, Pesaran, and Potter (1996), the generalized impulse response functions take the form

$$GI_Y = E[Y_{1+s}|\varepsilon_1 = .02, Y^0] - E[Y_{1+s}|Y^0] \text{ for } s \geq 0 \tag{2.31}$$

where Y is the vector of the variables and Y^0 is the history of Y up to time 0. The

---

$\gamma^+_{\hat{A}} = -.1079$.

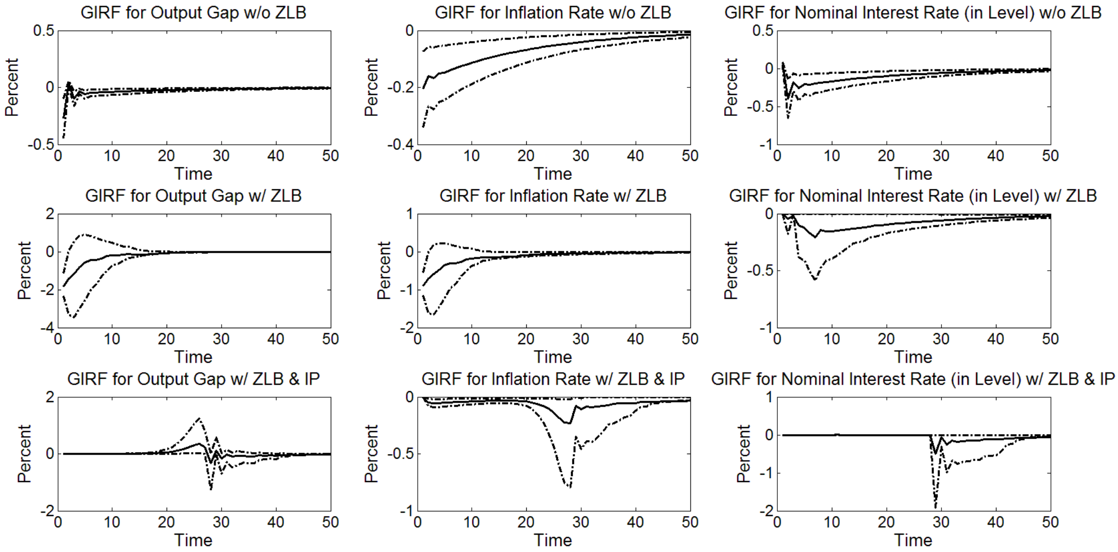
generalized impulse response functions are used to show how dynamics of the models in this paper are affected by the initial shock ε_1. The conditional expectations in (2.31) are approximated by Monte Carlo integration using the same data as above. Figure 2.3 plots the generalized impulse response functions of the output gap, the inflation rate, and the nominal interest rate (in level) with the unbroken lines and the associated 10th and 90th percentiles with the broken lines. The labeling conventions are identical to Figures 2.1 and 2.2. They largely confirm the observations based on Figures 2.1 and 2.2 above. For the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP"), the collapsing of the 90% confidence interval for the nominal interest rate into a line for long periods of time (the insignificant effect of the total factor productivity shock more generally as the 90% confidence interval contains zero throughout) suggests that the duration of the zero lower bound periods is longer for this model mainly because of the endogenous dynamics - that the private sector expects to continue to form expectations under uncertainty about the value of the current total factor productivity as long as the zero lower bound on the nominal interest rate is binding. What the generalized impulse response functions show is that the endogenous dynamics matters more than the current total factor productivity shock (as long as it keeps the economy at the zero lower bound for some periods) for the duration of the zero lower bound periods.

Finally, Figure 2.4 presents the 10th percentiles, the medians, and the 90th percentiles of the expectational errors for the current total factor productivity over time for the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP"). The errors take the form

$$e_t = E_t^p \hat{A}_t - \hat{A}_t. \tag{2.32}$$

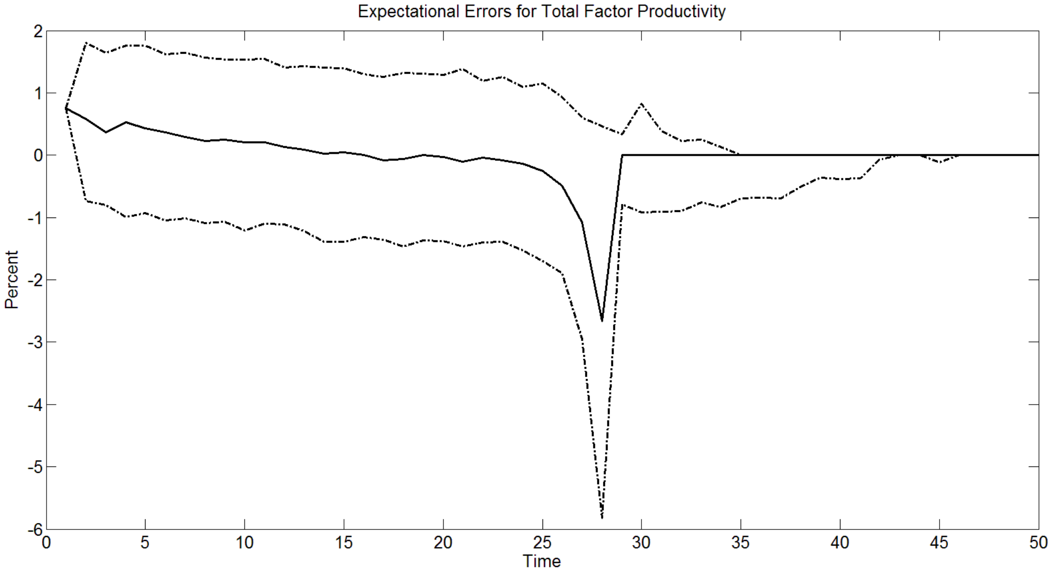
The unbroken line is for the medians and the broken lines are for the 10th and the 90th percentiles. The expectational errors for the standard new Keynesian model with and without the zero lower bound on the nominal interest rate ("With ZLB" and "Without ZLB" respectively) are zero because the private sector always observes the current total

Figure 2.3: Generalized Impulse Response Functions



The unbroken lines are for the mean differences given by (2.31) and the broken lines are for the 10th and the 90th percentiles of the differences. The labeling conventions are identical to Figures 2.1 and 2.2. $\phi_{\hat{A}} = .025$.

Figure 2.4: Expectational Errors at the Zero Lower Bound



The figure provides the distributions of the expectational errors for the current total factor productivity over time which arise due to the information problem at the zero lower bound. The unbroken line is for the median path of the errors and the broken lines are for the 10th and the 90th percentiles of the errors. $\phi_{\hat{A}} = .025$.

factor productivity in these models. Figure 2.4 shows that the expectational errors due to the information problem are small and insignificant (as the 90% confidence intervals contain zero). The median errors, which are positive at the beginning of the zero lower bound periods, turn negative along the way. The modesty of the errors (the time average of the median errors is -.0187%) suggests that the dynamics of "With ZLB & IP" in Figures 2.1, 2.2, and 2.3 are not driven by the expectational errors being large. Recall that there is at most one period delay in the observability of the current total factor productivity because the private sector can figure out what it is by the end of the period.[48] This implies that the past expectational errors do not affect the present and the future because their correct values are not known during the zero lower bound periods. Their effects stem from the dependence of the present and the future on the past through (2.25) as (a) the current expectational errors are fully reflected in the determination of the current endogenous variables and (b) the future endogenous variables depend on the current endogenous variables through (2.25).

The results so far provide some intuitions for the paths of the endogenous variables under "With ZLB & IP" in Figures 2.1 and 2.2 which corresponds to the new Keynesian model with the information problem at the zero lower bound. As shown by Figure 2.4, the (median) total factor productivities are overstated in the early parts of the zero lower bound periods. This implies that the economy is under more deflationary pressure than when the actual values of the total factor productivities are known because the inflation rate is negatively related to the total factor productivity[49]. Eventually, the (median) total factor productivities become understated and the economy faces increasing inflationary pressure. This inflationary pressure accelerates as old prices die out. Once the inflation rate recovers sufficiently, it kicks the nominal interest rate out of the zero lower bound. The initial overstatement and the eventual understatement of the total factor productivities are due to the sensitivity of the expected current total factor productivity to the inflation at the zero lower bound which was discussed

---

[48] Refer to subsubsection 2.3.2.2.

[49] For the new Keynesian models, the first order condition of the firms implies that the inflation rate can be expressed as a discounted sum of the firms' marginal costs (refer to (2.15)) which are negatively related to the total factor productivity.

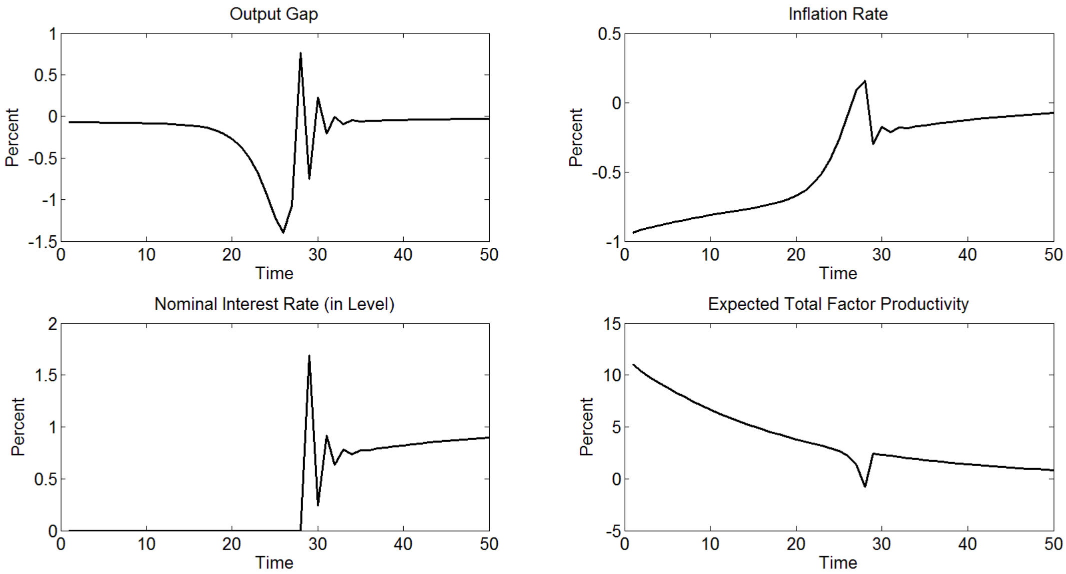
above. Because γ_x^+ and γ_π^+ in (2.25) are negative and the economy enters the zero lower bound from a state in which the output gap and the inflation rate are negative, their products sum up to be positive. Moreover, this overshoots the actual value of the total factor productivity because γ_π^+ is a large negative number (which stems from ϕ_π being larger than other ϕ's in (2.6) as explained above) and the inflation rate is also a large negative number in the neighborhood of the zero lower bound. This is the reason for the overstatement at the beginning of the zero lower bound periods. The eventual understatement begins once the inflation rate recovers sufficiently so that the product of γ_π^+ and the inflation rate is no longer big enough to overstate the total factor productivity. The recovery of the inflation rate is inevitable due to the mean-reverting nature of the total factor productivity as ρ in (2.5) is less than 1.

The figures above reveal that the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP") has very different dynamics compared to the standard new Keynesian model with and without the zero lower bound ("With ZLB" and "Without ZLB" respectively) which are similar apart from their initial reactions. Given that the standard models have been thoroughly studied in the literature, only the model with the information problem at the zero lower bound is analyzed further here. In order to take a closer look at the effect of the information problem, the model needs to be simulated under a particular sequence of the total factor productivity shocks that can mimic the median dynamics in Figure 2.1 as these are not associated with a particular sequence of the total factor productivity shocks but with the entire panel of the total factor productivity shocks. Working with a particular sequence of the total factor productivity shocks makes the computation of the expected paths of the variables at each point in time possible because it makes the history of these variables up to that point in time definite. Here, it is simulated under the shock sequence $(\varepsilon_1 = .02, 0, 0, \ldots, 0)$[50] which is how the ordinary impulse response functions are computed. This turns out to be sufficient for the purpose here as it captures the median dynamics under "With ZLB & IP" in Figure 2.1 satisfactorily. Figure 2.5 vali-

---

[50] This is also the value of ε_1 used above.

Figure 2.5: Impulse Response Functions



The figure provides the ordinary impulse response functions for the output gap, the inflation rate, the nominal interest rate, and the expected total factor productivity for the new Keynesian model with the information problem at the zero lower bound on the nominal interest rate ("With ZLB & IP" in Figures 2.1, 2.2, and 2.3).

dates this claim by plotting the ordinary impulse response functions. In what follows, the dynamics in Figure 2.5 are analyzed backward, starting from where the zero lower bound stops binding.

Why do the output gap, the inflation rate, and the nominal interest rate oscillate for a while once the zero lower bound periods end in Figure 2.5? Recall that $\hat{R}_t = \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_{\hat{A}} \hat{A}_t$ is the relevant interest rate rule when the zero lower bound on the nominal interest rate stops binding. Because $\phi_\pi > 1$, the nominal interest rate reacts more than one-to-one to the past inflation rate. Because there is a two-way feedback between the nominal interest rate and the output gap and the inflation rate, the strong reaction of the current nominal interest rate to the past inflation rate in turn implies the strong reaction of the current output gap and the current inflation rate to the current nominal interest rate. This is the reason for the oscillation of these variables.[51] Let us put this into the context of the question above. Upon exiting from

---

[51] Introducing an interest rate smoothing can weaken this oscillation. This does not alter the dynamics inside the zero lower bound qualitatively.

the zero lower bound, the nominal interest rate jumps up and this makes the output gap and the inflation rate jump down. This in turn decreases the nominal interest rate which increases the output gap and the inflation rate. This oscillation continues until being flattened out eventually.

In order to make sense out of what is going on during the zero lower bound periods in Figure 2.5, it is helpful to compute the expected paths of the relevant variables at various points of the zero lower bound periods. This is useful because the IS equation can be rewritten as (using the forward substitution)

$$\begin{aligned} E_t^p x_t &= -\frac{1}{\sigma}\sum_{s=1}^{\infty} E_t^p(\hat{R}_{t+s-1} - \pi_{t+s}) + \sum_{s=1}^{\infty} E_t^p u_{t+s-1} \\ &= -\frac{1}{\sigma}\sum_{s=1}^{\infty} E_t^p(\hat{R}_{t+s-1} - \pi_{t+s}) - \frac{1+\eta}{\sigma+\eta} E_t^p \hat{A}_t \qquad (2.33) \end{aligned}$$

where the second term in the second line follows because the functional form of u (refer to subsection 2.3.3) implies that u's beyond t in the second term in the first line cancel one another and the Phillips curve can be rewritten as (also using the forward substitution)

$$\pi_t = \lambda \sum_{s=0}^{\infty} \beta^s E_t^p x_{t+s} \qquad (2.34)$$

which involve the expectations of the current and future variables.[52] Using these representations, it is possible to dissect the expectations involved in the determination of the dynamics of the endogenous variables.

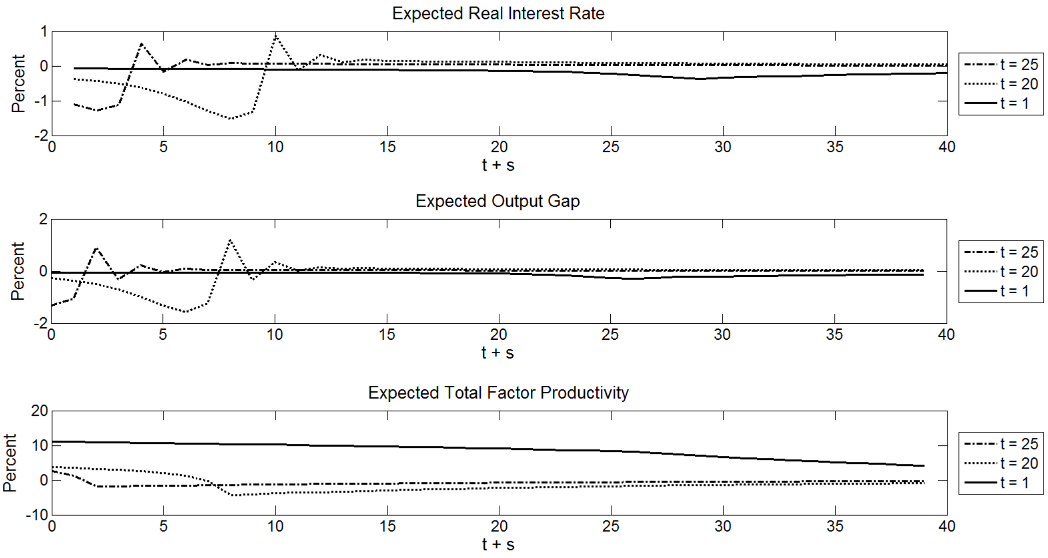
Figure 2.6 plots the expected paths of the relevant variables while the zero lower bound is binding. $t+s$ in the horizontal axis should be interpreted as s periods from t where t is the t_{th} period in the zero lower bound. The numbers along the horizontal axes are the values of s.

Why does the output gap start to decline and the inflation rate start to increase steeply as the end of the zero lower bound periods approaches in Figure 2.5? To answer this question, let us consider time periods 20 and 25 in Figure 2.5 which correspond

---

[52] See Gertler, Gali, and Clarida (1999) for a general discussion about the usefulness of forward-looking representations like these.

Figure 2.6: Expected Paths of Variables under Information Problem



The figure provides the expected paths of the real interest rate, the output gap, and the total factor productivity during the zero lower bound periods. $t+s$ in the horizontal axis should be interpreted as s periods from t where t corresponds to t_{th} period in the zero lower bound. The numbers along the horizontal axes are the values of s.

to $t = 20$ and 25 in Figure 2.6. The output gap takes a negative value in both because the first term of the second line in (2.33), which involves the sum of the expected real interest rates (row 1 in Figure 2.6) and is positive, is dominated by the second term of the second line in (2.33), which involves the current expected total factor productivity (the observation for $s = 0$ in row 3 of Figure 2.6) and is negative. It is also on its way down: the output gap in period 20 is not as large in magnitude as the one in period 25 because its first term is relatively larger due to the expected real interests from period 20 onward remaining negative longer than those from period 25 onward (row 1 in Figure 2.6). The inflation rate, which is represented by (2.34), takes a negative value because the positive expected output gaps in later periods are dominated by the negative expected output gaps in earlier periods (row 2 in Figure 2.6). It is also on its way up: the inflation rate in period 25 is higher than the one in period 20 because the expected output gaps from period 20 onward remain negative longer than those from period 25 onward (row 2 in Figure 2.6).

Why do the output gap and the inflation rate under the information problem react much less at the beginning of the zero lower bound periods? For this question, time period 1 in Figure 2.5 which corresponds to $t = 1$ in Figure 2.6 should be considered. The output gap is relatively smaller in period 0 because the second term in the second line of (2.33), which involves the current expected total factor productivity (the observation for $s = 0$ in row 3 of Figure 2.6) and is negative, is sufficiently offset by the first term in the second line of (2.33) which involves the expected real interest rates (row 1 in Figure 2.6) and is positive. This is because the expected real interest rates (row 1 in Figure 2.6) remain negative[53] for a long time from period 0 onward which makes the first term in the second line of (2.33) a considerably large positive number. The inflation rate, which is represented by (2.34), is also relatively higher in period 0 because the expected output gaps from period 0 onward are small (row 2 in Figure 2.6).

The analysis based on Figure 2.6 demonstrates how the information problem at the zero lower bound affects the dynamics of the output gap and the inflation rate through its effect on the expectations of the current and the future variables.

2.5.2 $\phi_{\hat{A}} = -.01$

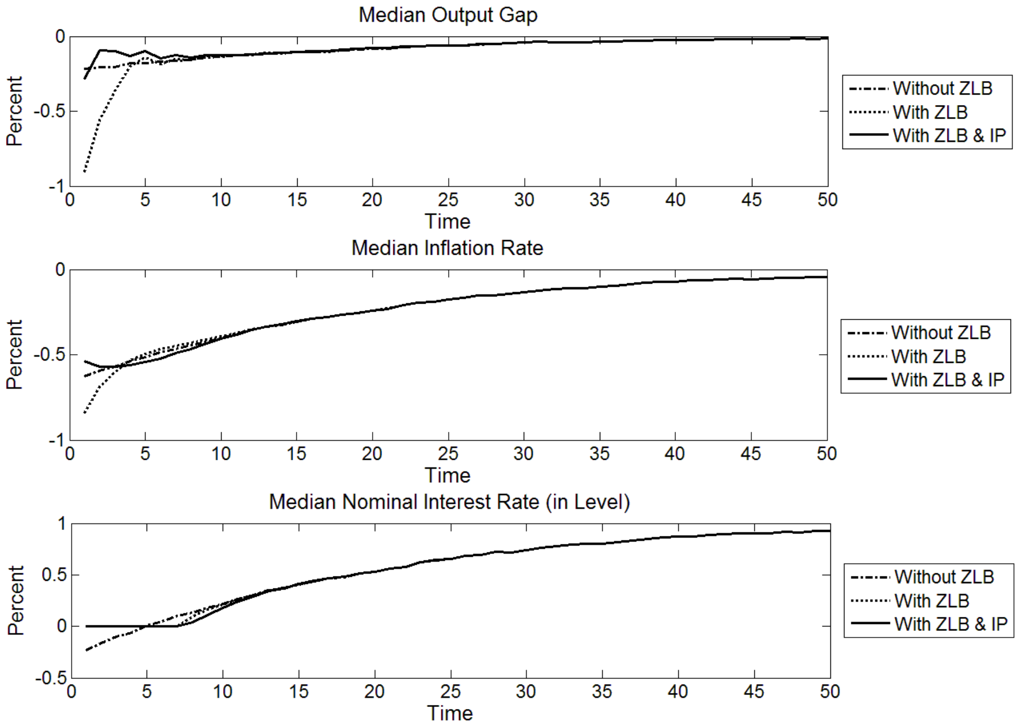
As in subsection 2.5.1, the standard new Keynesian model with and without the zero lower bound on the nominal interest rate ("With ZLB" and "Without ZLB" respectively) and the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP") will be analyzed using stochastic simulations. Some explanations will be omitted in order to avoid repeating subsection 2.5.1.

Figure 2.7 plots the median paths of the output gap, the inflation rate, and the nominal interest rate (in level) which were presented partially in Table 2.3. As shown above, (a) the duration of the zero lower bound periods is identical whether the private sector faces the information problem at the zero lower bound or not and (b) the deflationary pressure is more persistent with the information problem. Figure 2.8 provides the 10th and the 90th percentiles (broken lines) in addition to the medians (unbroken

---

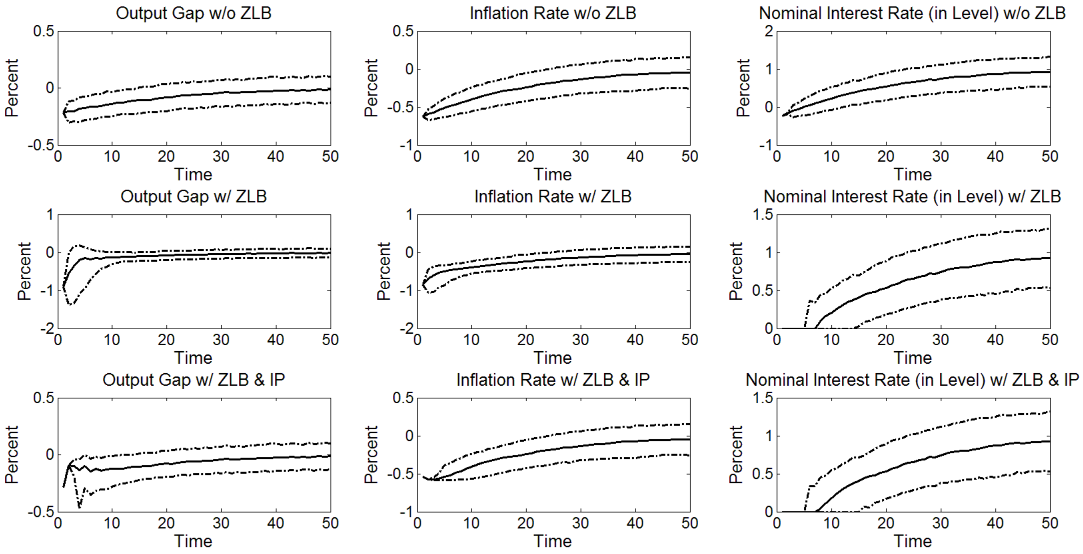
[53] The real interest rate is expressed as a deviation from its steady state value.

Figure 2.7: Median Paths of Endogenous Variables



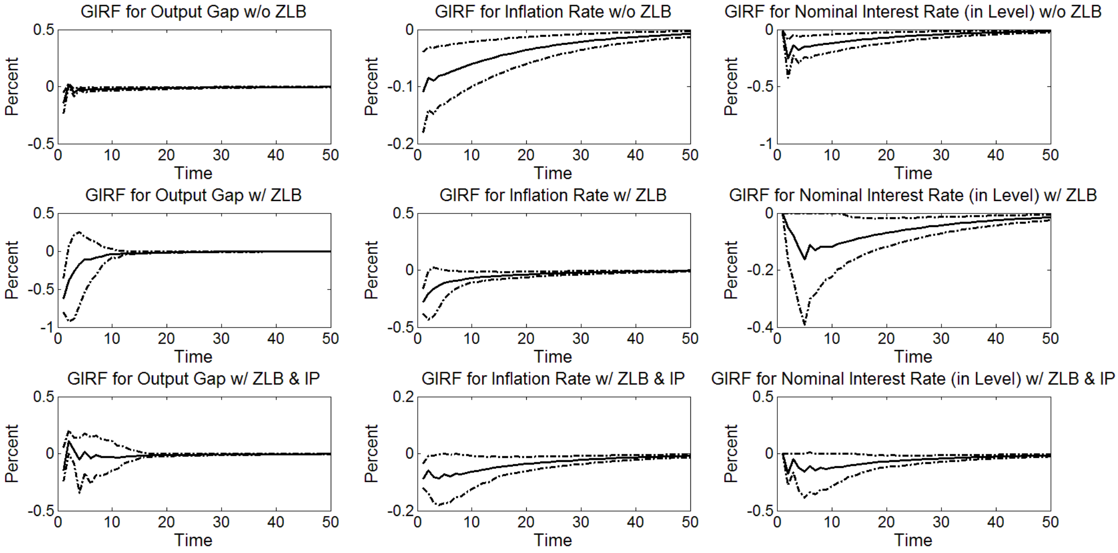
Without ZLB corresponds to the standard new Keynesian model without the zero lower bound on the nominal interest rate, With ZLB to the standard new Keynesian model with the zero lower bound, and With ZLB & IP to the new Keynesian model with the information problem at the zero lower bound. The medians are computed based on 1000 simulation rounds. $\phi_{\hat{A}} = -.01$.

Figure 2.8: Distributions of Endogenous Variables



The unbroken lines are for the medians of the variables and the broken lines are for the 10th and the 90th percentiles of the variables. The labeling conventions are identical to Figure 2.7. $\phi_{\hat{A}} = -.01$.

Figure 2.9: Generalized Impulse Response Functions


GIRF for Output Gap w/o ZLB
GIRF for Inflation Rate w/o ZLB
GIRF for Nominal Interest Rate (in Level) w/o ZLB
GIRF for Output Gap w/ ZLB
GIRF for Inflation Rate w/ ZLB
GIRF for Nominal Interest Rate (in Level) w/ ZLB
GIRF for Output Gap w/ ZLB & IP
GIRF for Inflation Rate w/ ZLB & IP
GIRF for Nominal Interest Rate (in Level) w/ ZLB & IP

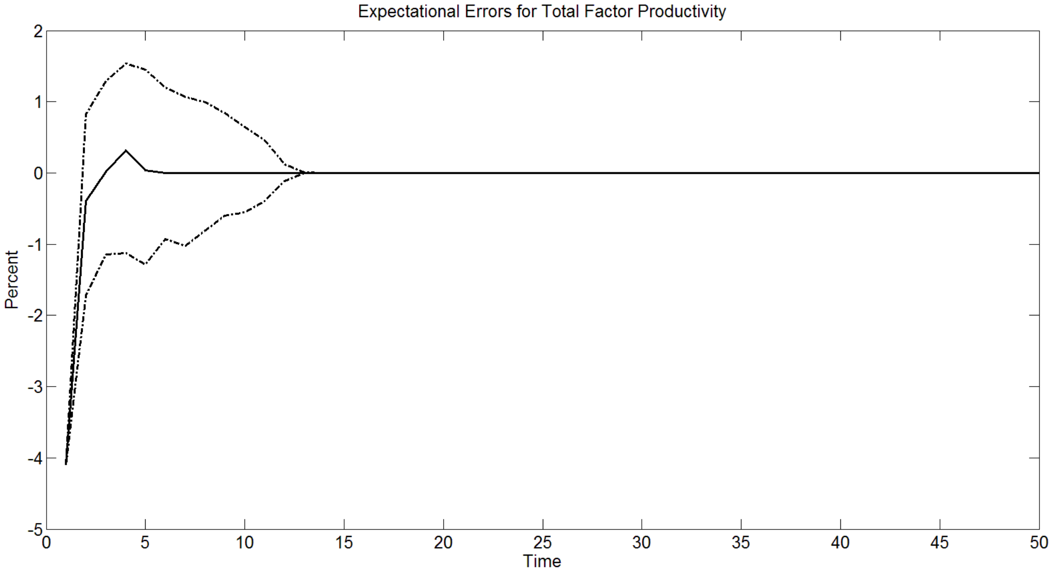
The unbroken lines are for the mean differences (in (2.31)) and the broken lines are for the 10th and the 90th percentiles of the differences. The labeling conventions are identical to Figures 2.7 and 2.8. $\phi_{\hat{A}} = -.01$.

lines). Unlike in Figure 2.2, the dynamics under the information problem ("With ZLB & IP") are not markedly different from those not under it ("With ZLB") in Figure 2.7. This is also confirmed by the generalized impulse response functions in Figure 2.9. In contrast with Figure 2.3, the 10th and the 90th percentiles of the nominal interest rate under "With ZLB & IP" no longer coincide for extended periods of time in Figure 2.9. This indicates that the mechanism at work here is relatively weaker in terms of its effect of on the duration of the zero lower bound periods.

As in Figure 2.4, taking a look at the expectational errors for the current total factor productivity is helpful for understanding the new Keynesian model with the information problem at the zero lower bound ("With ZLB & IP"). Figure 2.10 provides the median errors (unbroken line) together with the 10th and the 90th percentiles (broken lines). In contrast with those in Figure 2.4, the errors are negative first before becoming positive. The errors are also modest (the time average of the median errors is -.5890%) but larger than those for Figure 2.4.

The results so far are helpful for understanding the dynamics of the endogenous

Figure 2.10: Expectational Errors at the Zero Lower Bound



The unbroken line is for the median path of the errors and the broken lines are for the 10th and the 90th percentiles of the errors. $\phi_{\hat{A}} = -.01$.

variables under the information problem at the zero lower bound. The initial understatement of the total factor productivities implies that the economy is under less deflationary pressure than when the actual values of the total factor productivities are known. This is why the path of the inflation rate under the information problem ("With ZLB & IP") is U-shaped in Figure 2.7. Once the total factor productivities become overstated, the deflationary pressure kicks in but the extent of the overstatement is not large enough to bring about a strong and persistent delay in the recovery of the inflation rate. What is the reason for the initial understatement? Because γ_x^- and γ_π^- in (2.30) are positive and the economy enters the zero lower bound from a condition in which the output gap and the inflation rate are negative, their products sum up to be negative. Moreover, this undershoots the actual value of the total factor productivity because γ_π^- is a large positive number (because ϕ_π is larger than other ϕ's in (2.6) and this is reflected in (2.29)) and the inflation rate is a large negative number in the neighborhood of the zero lower bound. The eventual overstatement starts once the inflation rate recovers sufficiently so that the product of γ_π^- and the inflation rate is no longer large enough to put strong downward pressure.

The further analysis based on the impulse response functions is omitted as the explanations above are sufficient for the purpose here.

2.5.3 Discussion and Summary

In the previous subsections, the effect of the information problem at the zero lower bound on the nominal interest rate was analyzed. It was shown that the information problem can make the duration of the zero lower bound periods longer under positive $\phi_{\hat{A}}$'s. It was also shown that this is not always true under negative $\phi_{\hat{A}}$'s as the result gets overturned once $\phi_{\hat{A}}$ becomes sufficiently negative. Overall, on the both sides of zero for $\phi_{\hat{A}}$, there is a negative relationship between $\phi_{\hat{A}}$ and the extra periods inside the zero lower bound. As shown in section 2.2, the lower the value of $\phi_{\hat{A}}$, the more optimal the monetary policy is in the sense of bringing about lower volatilities of the output gap and the inflation rate. Thus, the monetary policy that is more optimal for the case without the information problem makes the zero lower bound periods shorter for the case with the information problem. In addition to its effect on the duration of the zero lower bound periods, the information problem makes the deflationary episode more persistent as well and this holds irrespective of the sign of $\phi_{\hat{A}}$. Appendix 2.8.2 provides figures for the median paths of the output gap, the inflation rate, and the nominal interest rate under other values of $\phi_{\hat{A}}$ in order to demonstrate that the analysis in subsections 2.5.1 and 2.5.2 are qualitatively robust to using different values of $\phi_{\hat{A}}$.

Having seen that the behavior of the new Keynesian model with the information problem at the zero lower bound depends on the sign of $\phi_{\hat{A}}$, one may wonder how the frequency of the zero lower bound periods depends on the sign of $\phi_{\hat{A}}$. Table 2.5 provides the average and the maximum frequency of the zero lower bound periods for the values of $\phi_{\hat{A}}$ considered above. The conditions under which these are generated are identical to the ones for Table 2.1 in section 2.2 (e.g., 1000 simulation rounds, 2000 periods each after discarding the initial 200 periods). Table 2.5 shows that the average and the maximum frequency of the zero lower bounds (given by "Ave Freq" and "Max Freq" respectively) are increasing in $\phi_{\hat{A}}$. Given that the frequency of the zero lower

Table 2.5: Frequency of the Zero Lower Bound Periods

	$\phi_{\hat{A}} = .025$	$\phi_{\hat{A}} = .02$	$\phi_{\hat{A}} = .015$	$\phi_{\hat{A}} = .01$	$\phi_{\hat{A}} = .005$	$\phi_{\hat{A}} = -.005$	$\phi_{\hat{A}} = -.01$	$\phi_{\hat{A}} = -.015$
Ave Freq	38.91	24.74	16.54	11.41	8.13	4.39	3.32	2.61
Max Freq	275	144	102	74	57	29	20	13

Ave Freq and Max Freq are the average and the maxium frequency of the zero lower bound periods respectively.

bound periods in the US and the European data belongs to a 5 to 10% interval (an interval rather than a point estimate is used here as the latter is sensitive to which year is chosen as the initial year in the data set), one may ask whether the new Keynesian model with the information problem at the zero lower bound is consistent with this interval. It is obvious that the negative values of $\phi_{\hat{A}}$ in Table 2.5 are not consistent with it as the average and the maximum frequency are too low in these cases. Among the positive $\phi_{\hat{A}}$'s in Table 2.5, $\phi_{\hat{A}} = .025$ is consistent with it as the maximum frequency of the zero lower bound periods obtained from the simulation, which is around 14% (275 out of 2000 periods), contains the 5 to 10% interval. Figure 2.11 gives a histogram of frequencies of the zero lower bound periods for $\phi_{\hat{A}} = .025$. It shows that the 5 to 10% interval, which corresponds to between 100 and 200 periods in the horizontal axis of the histogram, occurs with a non-negligible frequency of around 4.2%.[54] This is an improvement relative to the standard new Keynesian model with $\phi_{\hat{A}} = 0$ (a commonly used model without the information problem) for which this frequency is zero.

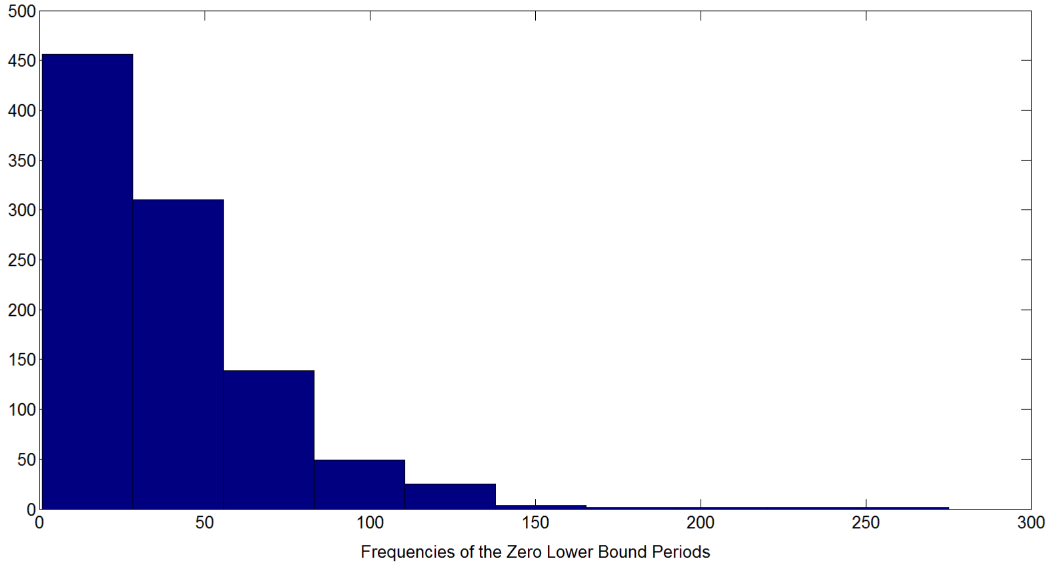
2.6 Policy Analysis

It was discussed in section 2.2 that the loss associated with monetary policy is increasing in $\phi_{\hat{A}}$ in the absence of the information problem at the zero lower bound. What this section does is essentially comparing the loss under the information problem with the loss not under it for various values of $\phi_{\hat{A}}$. The information problem associated with the current total factor productivity at the zero lower bound on the nominal interest rate can be resolved if the central bank announces what its value is while the

---

[54] Given that there is only one source of shocks (the Total Factor Productivity shocks) in this paper, this frequency should not be considered low.

Figure 2.11: Histogram of Frequencies of the Zero Lower Bound Periods for $\phi_{\hat{A}} = .025$



zero lower bound is binding. In terms of the models in section 2.5, the standard new Keynesian model with the zero lower bound ("With ZLB") corresponds to the case in which the central bank makes this announcement and the new Keynesian model with the information problem ("With ZLB & IP") to the case in which it does not. In order to assess the benefits of the increased transparency from the announcement, a periodic loss function of the form

$$L_t = \frac{1}{2}(\alpha x_t^2 + \pi_t^2) \tag{2.35}$$

is used to compare these models under various values of α. (2.35) is a quadratic loss function which is commonly used as the central bank's objective function to be minimized in the literature. For the purpose here, (2.35) can be thought of as a summary measure for the volatilities of the output gap and the inflation rate with α being the relative weight that the central bank attaches to the output gap over the inflation rate. In many mainstream models, α tends to be a small number. For instance, Woodford (2003)'s calibration of the standard new Keynesian model to the US data yields $\alpha = .003$. In a related context, Soderstrom, Soderlind, and Vredin (2002) show that α should not exceed .1 for their new Keynesian model to match the behavior of the US data.[55] In what follows, for each value of $\phi_{\hat{A}}$ considered in section 2.5, the time average of (2.35)[56] is computed based on the median paths of the output gap and the inflation rate using a grid of values for α that range from 0 to .1. The median paths used here correspond to those partially shown in Tables 2.3 and 2.4 in section 2.5 (and shown fully in Figures 2.1, 2.7, and the figures in Appendix 2.8.2). Using a finite time average of (2.35) is sufficient as a loss function in the sense that the output gap and the inflation rate under "With ZLB" and "With ZLB & IP" converge to the same paths quickly in these figures. The output gap and the inflation rate from the point of convergence onward are irrelevant for comparing losses of these models because this amounts to adding the same number to them.

Table 2.6 provides ratios of the loss under "With ZLB & IP" to the loss under "With ZLB." The ratio measures the loss due to the information problem at the zero lower bound whose source is the lack of explicit communication by the central bank about the value of the current total factor productivity. The table shows that whereas the information problem brings about larger losses for the positive $\phi_{\hat{A}}$'s, the opposite is true for the negative $\phi_{\hat{A}}$'s. As shown in Figures 2.1 and 2.7 (and the figures in Appendix 2.8.2), even though the output gap and the inflation rate react more initially under "With ZLB," the deflationary pressure is more persistent under "With ZLB & IP." The outcomes under different signs of $\phi_{\hat{A}}$ differ due to the shifts in the relative magnitudes of these two forces.

One can think of the increased central bank transparency in the form of information revelation as a type of forward guidance. Forward guidance, which is currently in use as a policy tool by some leading central banks[57], involves communication about the future course of monetary policy in order to shape the public's expectations about the future. In the model, the explicit communication about the value of the current total factor productivity by the central bank allows the private sector to form correct

---

[55] Soderstrom, Soderlind, and Vredin (2005) estimate this model and show that α is zero.

[56] This is isomorphic to the discounted sum of (2.35) with the discount factor of 1.

[57] See VOX articles by Peter Praet (http://www.voxeu.org/article/forward-guidance-and-ecb), which explains the European Central Bank's practice, and Spencer Dale and James Talbot (http://www.voxeu.org/article/forward-guidance-uk), which discusses the Bank of England's practice.

Table 2.6: Loss Due to Information Problem at the Zero Lower Bound

	$\alpha = 0$	$\alpha = .02$	$\alpha = .04$	$\alpha = .06$	$\alpha = .08$	$\alpha = .1$
$\phi_{\hat{A}} = .025$	1.31	1.29	1.28	1.27	1.25	1.24
$\phi_{\hat{A}} = .02$	1.08	1.07	1.06	1.05	1.04	1.03
$\phi_{\hat{A}} = .015$	1.05	1.04	1.04	1.03	1.03	1.02
$\phi_{\hat{A}} = .01$	1.05	1.04	1.04	1.04	1.04	1.03
$\phi_{\hat{A}} = .005$	1.04	1.04	1.04	1.04	1.04	1.04
$\phi_{\hat{A}} = -.005$	.80	.79	.78	.77	.77	.76
$\phi_{\hat{A}} = -.01$	.92	.91	.90	.90	.90	.89
$\phi_{\hat{A}} = -.015$	.98	.98	.98	.97	.97	.97

The table gives ratios of the loss under "With ZLB & IP" to the loss under "With ZLB" for the values of $\phi_{\hat{A}}$ considered in section 2.5. α is the relative weight that the central bank attaches to the output gap over the inflation rate in (2.35).

expectations about the future when the information flow from the nominal interest rate is impeded by the zero lower bound. In this sense, the information revelation by the central bank can be considered as forward guidance. This type of forward guidance is called "Delphic" (Campbell, Evans, Fisher, and Justiniano, 2012) because it is concerned with the transmission of information and involves forecasts of the future. Table 2.6 shows that while the increased central bank transparency is beneficial when $\phi_{\hat{A}}$ is positive, the opposite is true when $\phi_{\hat{A}}$ is negative.

2.7 Conclusion

This paper recognizes that the interest rate contains valuable information about the state of the economy, and this information flow is impeded when the nominal interest rate hits the zero lower bound. This impediment to the information flow at the zero lower bound brings about a discrete jump in uncertainty for the private sector which is referred to as the information problem at the zero lower bound. The paper shows that under some parametrizations, the information problem at the zero lower bound makes (a) the duration of zero lower bound periods longer and (b) the deflationary pressure in and out of the zero lower bound more persistent. Moreover, it makes the exit from the zero lower bound disorderly by bringing about large fluctuations in the output gap, the inflation rate, and the nominal interest rate. The policy analysis that builds on these observations suggests that increased central bank transparency in the form of information revelation is beneficial. This information revelation is a form of forward guidance by which the central bank communicates the expected course of monetary policy in order to manage the private sector's expectations about the future. Thus, this paper provides a theoretical support for forward guidance which is currently practiced by a number of leading central banks. However, there exist parametrizations for which the increased central bank is not beneficial and this is interesting because the model in this paper does not have features that are especially conducive for generating this result.

2.8 Appendix

2.8.1 Derivation of (2.3)

Recall that a function of the form

$$\sum_{k=-1}^{1} \phi_{x,k} E_t x_{t-k} + \sum_{k=-1}^{1} \phi_{\pi,k} E_t \pi_{t-k}$$

is being considered here. The unconstrained linear rational expectations solutions of the standard new Keynesian model take the form

$$x_t = \psi^x_{x,0} x_{t-1} + \psi^x_{\pi,0} \pi_{t-1} + \psi^x_{\hat{A},0} \hat{A}_t,$$

$$E_t x_{t+1} = \psi^x_{x,-1} x_{t-1} + \psi^x_{\pi,-1} \pi_{t-1} + \psi^x_{\hat{A},-1} \hat{A}_t,$$

$$\pi_t = \psi^\pi_{x,0} x_{t-1} + \psi^\pi_{\pi,0} \pi_{t-1} + \psi^\pi_{\hat{A},0} \hat{A}_t, \text{ and}$$

$$E_t \pi_{t+1} = \psi^\pi_{x,-1} x_{t-1} + \psi^\pi_{\pi,-1} \pi_{t-1} + \psi^\pi_{\hat{A},-1} \hat{A}_t$$

when the nominal interest rate is set according to the function above. Substituting these into $\sum_{k=-1}^{1}\phi_{x,k}E_t x_{t-k}+\sum_{k=-1}^{1}\phi_{\pi,k}E_t\pi_{t-k}$ and collecting the like terms yields

$$(\phi_{x,1}+\phi_{x,0}\psi_{x,0}^{x}+\phi_{\pi,0}\psi_{x,0}^{\pi}+\phi_{x,-1}\psi_{x,-1}^{x}+\phi_{\pi,-1}\psi_{x,-1}^{\pi})x_{t-1}$$

$$+(\phi_{\pi,1}+\phi_{x,0}\psi_{\pi,0}^{x}+\phi_{\pi,0}\psi_{\pi,0}^{\pi}+\phi_{x,-1}\psi_{\pi,-1}^{x}+\phi_{\pi,-1}\psi_{\pi,-1}^{\pi})\pi_{t-1}$$

$$+(\phi_{x,0}\psi_{\hat{A},0}^{x}+\phi_{\pi,0}\psi_{\hat{A},0}^{\pi}+\phi_{x,-1}\psi_{\hat{A},-1}^{x}+\phi_{\pi,-1}\psi_{\hat{A},-1}^{\pi})\hat{A}_t.$$

By setting

$$\phi_x=\phi_{x,1}+\phi_{x,0}\psi_{x,0}^{x}+\phi_{\pi,0}\psi_{x,0}^{\pi}+\phi_{x,-1}\psi_{x,-1}^{x}+\phi_{\pi,-1}\psi_{x,-1}^{\pi},$$

$$\phi_\pi=\phi_{\pi,1}+\phi_{x,0}\psi_{\pi,0}^{x}+\phi_{\pi,0}\psi_{\pi,0}^{\pi}+\phi_{x,-1}\psi_{\pi,-1}^{x}+\phi_{\pi,-1}\psi_{\pi,-1}^{\pi},\text{ and}$$

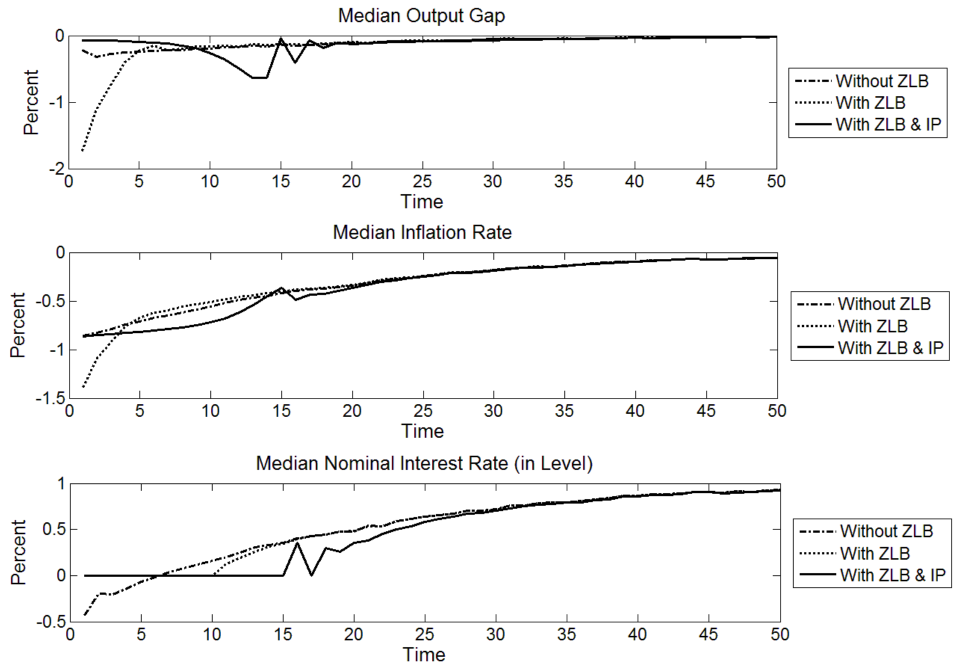
$$\phi_{\hat{A}}=\phi_{x,0}\psi_{\hat{A},0}^{x}+\phi_{\pi,0}\psi_{\hat{A},0}^{\pi}+\phi_{x,-1}\psi_{\hat{A},-1}^{x}+\phi_{\pi,-1}\psi_{\hat{A},-1}^{\pi},$$

the expression above reduces to

$$\phi_x x_{t-1}+\phi_\pi\pi_{t-1}+\phi_{\hat{A}}\hat{A}_t.$$

2.8.2 Median Paths of Endogenous Variables under Various Values of $\phi_{\hat{A}}$

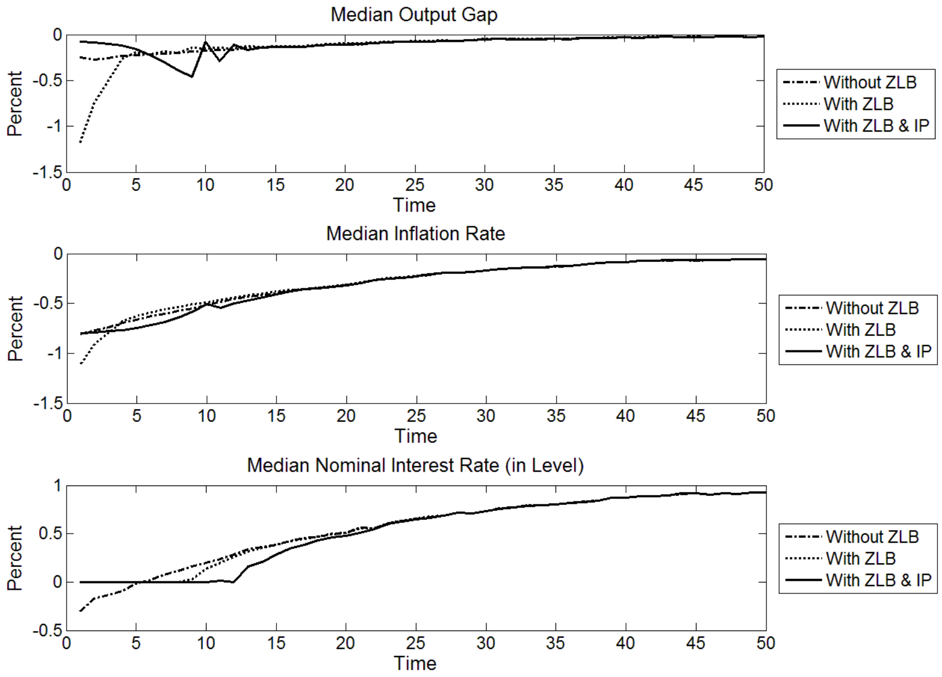
Figure 2.12: $\phi_{\hat{A}} = .02$


Median Output Gap

Median Inflation Rate

Median Nominal Interest Rate (in Level)

Figure 2.13: $\phi_{\hat{A}} = .015$


Median Output Gap

Median Inflation Rate

Median Nominal Interest Rate (in Level)

Figure 2.14: $\phi_{\hat{A}} = .01$

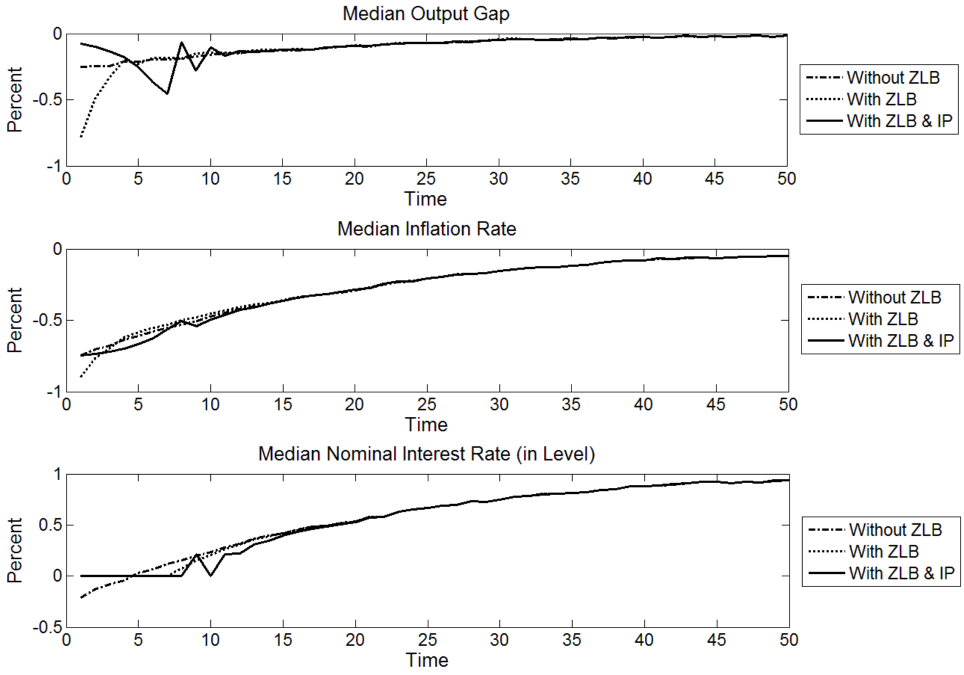


Figure 2.15: $\phi_{\hat{A}} = .005$

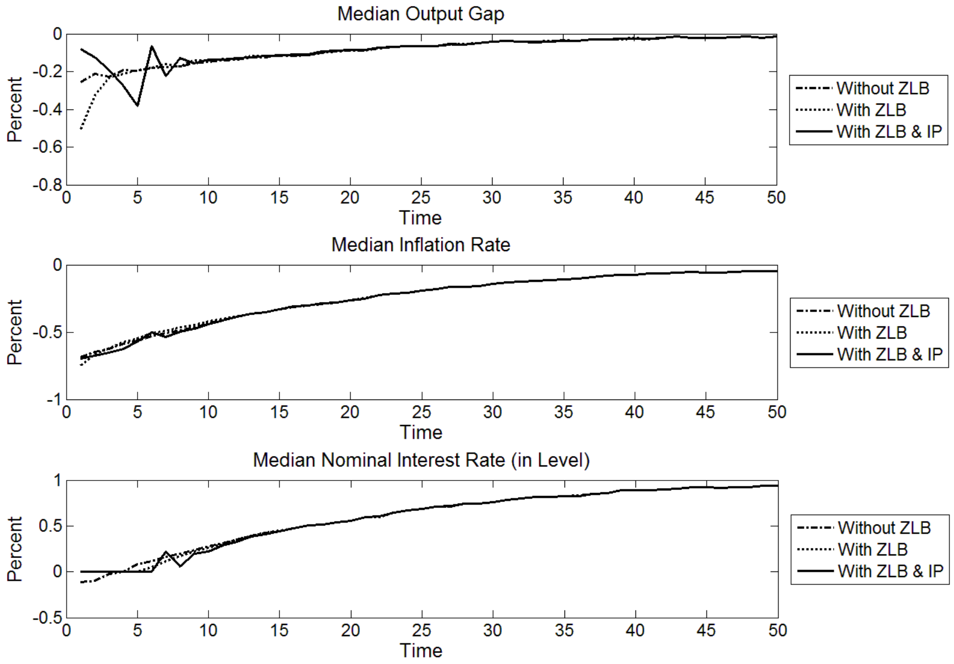


Figure 2.16: $\phi_{\hat{A}} = -.005$

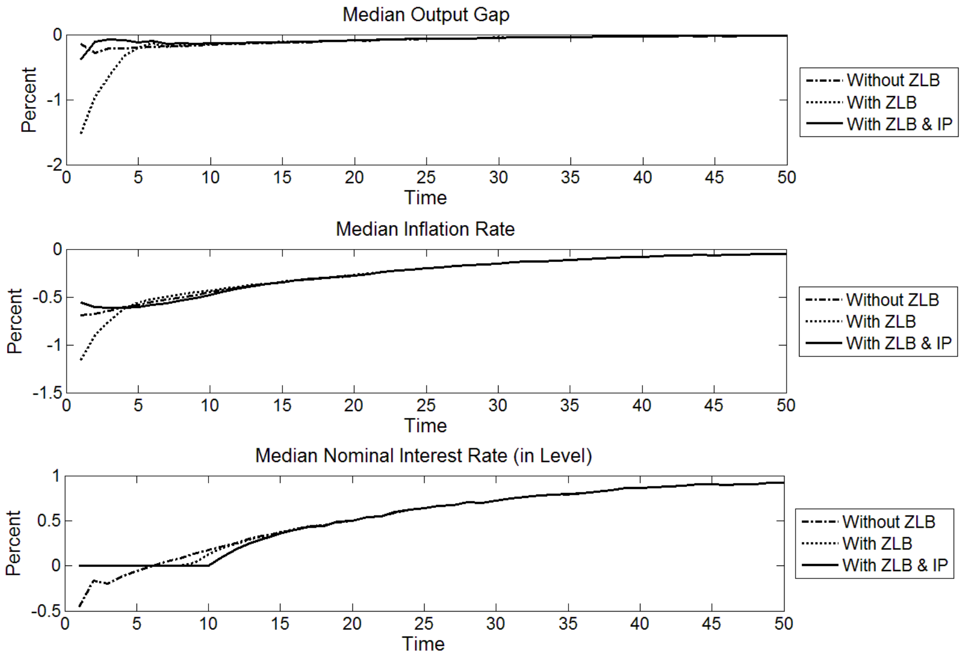
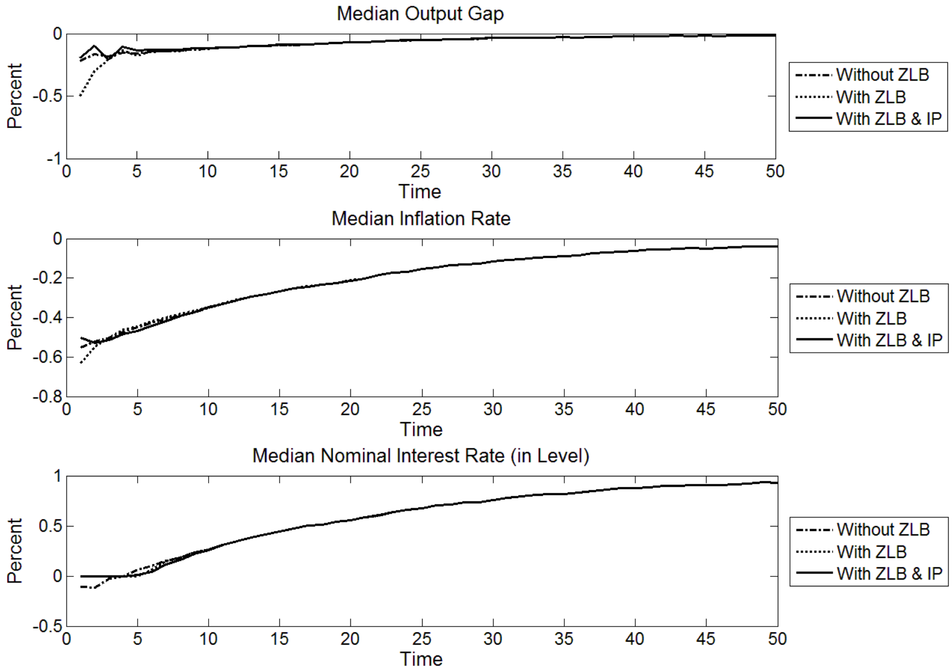


Figure 2.17: $\phi_{\hat{A}} = -.015$



2.9 References

1. Adam, K. and Billi, R. M. (2007). "Discretionary monetary policy and the zero lower bound on nominal interest rates," Journal of Monetary Economics, Elsevier, vol. 54(3), pages 728-752, April.

2. Álvarez, L. J., Dhyne E., Hoeberichts, M., Kwapil, C., Le Bihan, H., Lünnemann, P., Martins, F., Sabbatini, R., Stahl, H., Vermeulen, P., and Vilmunen, J. (2006). "Sticky prices in the Euro area: a summary of new micro-evidence," Journal of the European Economic Association, MIT Press, vol. 4(2-3), pages 575-584.

3. Aruoba, S. B. and Schorfeide, F. (2012). "Macroeconomic dynamics near the ZLB: a tale of two equilibria," Working Paper.

4. Bianchi, F. and Melsoi, L. (2014). "Escaping the great recession," Working Paper.

5. Blinder, A. S. (1998). "Central banking in theory and practice," MIT Press.

6. Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., and Jansen, D. (2008). "Central bank communication and monetary policy: a survey of theory and evidence," Journal of Economic Literature, American Economic Association, vol. 46(4), pages 910-45, December.

7. Bodenstein, M., Erceg, C. J., and Guerrieri, L. (2009). "The effects of foreign shocks when interest rates are at zero," International Finance Discussion Papers 983, Board of Governors of the Federal Reserve System (U.S.).

8. Calvo, G. A. (1983). "Staggered prices in a utility-maximizing framework," Journal of Monetary Economics, Elsevier, vol. 12(3), pages 383-398, September.

9. Cameron, A. C. and Trivedi, P. K. (2005). "Microeconometrics: methods and applications," Cambridge University Press.

10. Campbell, J. R., Evans, C. L., Fisher, J. D. M., and Justiniano, A. (2012). "Macroeconomic effects of Federal Reserve forward guidance," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 44(1), pages 1-80.

11. Christiano, L. J., Eichenbaum, M., and Rebelo, S. (2011). "When is the government spending multiplier large?," Journal of Political Economy, University of Chicago Press, vol. 119(1), pages 78-121.

12. Christiano, L. J., Trabandt, M., and Walentin, K. (2010). "DSGE models for monetary policy analysis," Handbook of Monetary Economics, in Friedman, B. M. and Woodford, M. (ed.), Handbook of Monetary Economics, Elsevier, vol. 3, chapter 7, pages 285-367.

13. Cooley, T. F. and Prescott, E. C. (1995). "Economic growth and business cycles," Frontiers of Business Cycle Research, in Cooley, T. F. (ed.), Frontiers of Business Cycle Research, Princeton Univeristy Press, chapter 1, pages 1-38.

14. Cukierman, A. (2009). “The Limits of Transparency,” Economic Notes, Banca Monte dei Paschi di Siena SpA, vol. 38(1-2), pages 1-37.

15. Dixit, A. K. and Stiglitz, J. E. (1977). “Monopolistic competition and optimum product diversity,” American Economic Review, American Economic Association, vol. 67(3), pages 297-308, June.

16. Eggertsson, G. B. and Woodford, M. (2003). “The zero bound on interest rates and optimal monetary policy,” Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 34(1), pages 139-235.

17. Fernández-Villaverde, J., Gordon, G., Guerrón-Quintana, P., and Rubio-Ramírez, J. F. (2012). “Nonlinear adventures at the zero lower bound,” Working Papers 12-10, Federal Reserve Bank of Philadelphia.

18. Gali, J. (2008). “Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework,” The Princeton University Press.

19. Gertler, M., Gali, J., and Clarida, R. (1999). “The science of monetary policy: a new Keynesian perspective,” Journal of Economic Literature, American Economic Association, vol. 37(4), pages 1661-1707, December.

20. Gosselin, P., Lotz, A., and Wyplosz, C. (2009). “Interest rate signals and central bank transparency,” NBER International Seminar on Macroeconomics 2007, National Bureau of Economic Research, Inc., pages 9-51.

21. Guerrieri, L. and Iacoviello, M. (2014). “OccBin: a toolkit for solving dynamic models with occasionally binding constraints easily,” Working Paper, Federal Reserve Board.

22. Gürkaynak, R. S., Sack, B., and Swanson, E. (2005). “Do actions speak louder than words? The response of asset prices to monetary policy actions and statements,” International Journal of Central Banking, International Journal of Central Banking, vol. 1(1), May.

23. Holden, T. and Paetz, M. (2012). “Efficient simulation of DSGE models with inequality constraints,” School of Economics Discussion Papers 1612, School of Economics, University of Surrey.

24. Hubert, P. (2013). “Do central bank forecasts influence private agents? Forecasting performance vs. signals,” Working Paper, OFCE - Sciences Po.

25. Judd, K. L., Maliar, L., and Maliar, S. (2011). “A cluster-grid projection method: solving problems with high dimensionality,” Working Paper.

26. Jung, T., Teranishi, Y., and Watanabe, T. (2005). “Optimal monetary policy at the zero-interest-rate bound,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 37(5), pages 813-35, October.

27. Koop, G., Pesaran, M. H., and Potter, S. M. (1996). “Impulse response analysis in nonlinear multivariate models,” Journal of Econometrics, Elsevier, vol. 74(1), pages 119-147, September.

28. Kuttner, K. N. (2001). “Monetary policy surprises and interest rates: evidence from the Fed funds futures market,” Journal of Monetary Economics, Elsevier, vol. 47(3), pages 523-544, June.

29. Nakata, T. (2012). “Uncertainty at the zero lower bound,” Working Paper.

30. Nakov, A. (2008). "Optimal and simple monetary policy rules with zero floor on the nominal interest rate," International Journal of Central Banking, International Journal of Central Banking, vol. 4(2), pages 73-127, June.

31. Peek, J., Rosengren, E. S., and Tootell, G. M. B. (1999). “Is bank supervision central to central banking?,” The Quarterly Journal of Economics, MIT Press, vol. 114(2), pages 629-653, May.

32. Romer, D. H. and Romer, C. D. (2000). “Federal reserve information and the behavior of interest tates,” American Economic Review, American Economic Association, vol. 90(3), pages 429-457, June.

33. Sims, C. A. (2002). “The role of models and probabilities in the monetary policy process,” Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 33(2), pages 1-62.

34. Söderström, U., Söderlind, P., and Vredin, A. (2002). "New-Keynesian models and monetary policy: a reexamination of the stylized facts," Working Paper Series in Economics and Finance 511, Stockholm School of Economics.

35. _______ (2005). “New-Keynesian models and monetary policy: a re-examination of the stylized facts,” Scandinavian Journal of Economics, Wiley Blackwell, vol. 107(3), pages 521-546, September.

36. Walsh, C. E. (2003). “Monetary theory and policy, 2nd edition,” MIT Press Books, MIT Press.

37. Werning, I. (2012). “Managing a liquidity trap: monetary and fiscal policy,” Working Paper, MIT.

38. Woodford, M. (2005). “Central bank communication and policy effectiveness,” Proceedings - Economic Policy Symposium - Jackson Hole, Federal Reserve Bank of Kansas City, pages 399-474, August.

39. _______ (2011). “Simple analytics of the government expenditure multiplier,” American Economic Journal: Macroeconomics, American Economic Association, vol. 3(1), pages 1-35, January.

Chapter 3

Growth under Knightian Uncertainty: Application of Interactive Trial and Error Learning

Abstract

Knightian Uncertainty is defined as a situation in which the structure of the economic system is unknown to its agents. This implies that the agents face difficulty with forming expectations about the future which are necessary for deriving their optimal investment rules. To cope, they follow Interactive Trial and Error Learning ($ITEL$) (Young, 2009) as behavioral rules for choosing investment portfolios. Under $ITEL$, the agents occasionally experiment with new actions to improve their payoffs even when they are satisfied with their current actions. This chapter presents an application of $ITEL$ which is concerned with the growth of aggregate variables under Knightian Uncertainty. The growth model can match several business cycle features of the US real aggregate wealth data such as the durations and the amplitudes of contractions and expansions.

3.1 Introduction

This paper is concerned with the growth of aggregate variables under Knightian Uncertainty. Knight, in his book *Risk, Uncertainty, and Profit* (1921), distinguished "uncertainty" from "risk" as follows:

> "Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated. The term "risk," as loosely used in everyday speech and in economic discussion, really covers two things which ... are categorically different. ... The essential fact is that "risk" means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character. ... It will appear that a measurable uncertainty, or "risk" proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. We shall accordingly restrict the term "uncertainty" to cases of the non-quantitative type."

Knight (1921) was not alone in holding this view: Keynes (1921, 1937) and Mises (1940, 1949) held similar views about how uncertainty is distinguished from risk.[1] Roughly speaking, Knightian Uncertainty is uncertainty whose structure is not well-known, in contrast to risk whose structure is well-known. With the global economy entering a severe recession in 2008, the concept of Knightian Uncertainty became the center of attention among mainstream macroeconomists. For instance, Blanchard, in his *Economist* article, ascribed the evolution of the financial crisis to Knightian Uncertainty and urged that removing this uncertainty is essential for recovery.[2,3]

This paper follows this recently rekindled interest in the subject by incorporating Knightian Uncertainty into the model. The model agents face Knightian Uncertainty because the structure of the economic system, in particular the mechanism that generates the return on investment, is unknown to them. This implies that it is not obvious to the agents in charge of choosing investment portfolios how they should form their expectations about the future because they do not have any knowledge about the relevant process upon which their expectations should be conditioned. It follows that

---

[1] See Hauwe (2011) and Eric Schliesser's blog entry at http://www.newappsblog.com/2012/03/weekly-philo-of-economics-keynes-and-knightian-uncertainbty.html.

[2] See http://www.economist.com/node/13021961.

[3] See Ricardo Caballero's VOX blog entries for an overview: http://www.voxeu.org/article/crisis-and-how-fix-it-part-1-causes and http://www.voxeu.org/article/crisis-and-how-fix-it-part-2-solutions.

the agents cannot derive their optimal portfolio choice rules as these depend on their expectations about the future in a dynamic setting. In order to address this issue, this paper introduces Interactive Trial and Error Learning ($ITEL$) (Young, 2009; Pradelski and Young, 2012) as behavioral rules for portfolio choice. Under $ITEL$, the agents search for actions that lead to higher payoffs by trial and error. This involves occasional experiments with new actions even when they are satisfied with their current actions.[4] Because this adaptive learning rule is completely uncoupled, that is it depends only on one's own action and own payoff, it is compatible with the model environment in which an agent in one location is constrained to know nothing about agents in other locations.[5] $ITEL$, which consists of fairly intuitive learning rules, is widely applicable because it is not based on an explicit model. More importantly, under certain conditions, it can implement Pure Nash Equilibria (PNEs) of the relevant investment game a very high proportion of the time.[6,7] Thus, its use can be justified on the grounds that it is fairly intuitive, widely applicable, and implements an outcome consistent with rationality (i.e., PNEs) very frequently.

As mentioned above, this paper deals with the growth of aggregate variables. Here, dynamic equations describing the evolution of the aggregate wealth and the aggregate consumption are derived by imposing certain restrictions on the beliefs of the agents in charge of making consumption and saving decisions. It will be shown that these restrictions on the beliefs transform the optimal control problem under Knightian Uncertainty into the optimal control problem under imagined or virtual certainty.

---

[4] Bandit problems are another example in which model agents face the trade-off between an experimentation with a new action and an exploitation of an action that is believed to deliver the best payoff. See Bergemann and Valimaki (2006) for a short survey of the literature.

[5] Another type of adaptive learning rule that is compatible with this information structure is Reinforcement Learning (see Harley 1981; Roth and Erev, 1995; Camerer and Ho, 1999). For empirical evidence about this learning rule, see Erev and Roth (1998) and Song, Jang, Hanssens, and Suh (2014). For its adaptability to macroeconomics and finance, see Choi, Laibson, Madrian, and Matrick (2009) and Anufriev, Assenza, Hommes, and Massaro (2013).

[6] This paper is not concerned with Mixed Nash Equilibria.

[7] Fudenberg and Levine (1998) provides a survey of the literature about the convergence to Nash Equilibrium under "fictitious play." The implementation of this learning rule requires more information than $ITEL$.

In contrast to most growth models in the literature which are concerned with the long-run balanced-growth of aggregate variables[8], the model in this paper analyzes both balanced-growth and unbalanced-growth of aggregate variables at every point in time. That is, the model here is concerned with both trend and cycles of the aggregate variables. It will be shown that whereas the balanced-growth is usually associated with the PNEs of the relevant investment game, the unbalanced-growth is usually associated with the disequilibria of the same game. Because $ITEL$ selects the PNEs of the investment game a high proportion of the time, it follows that the fluctuations in the growth rates of the aggregate variables are usually associated with the disequilibria of the investment game. And these fluctuations are recurrent over time.

The model in this paper is calibrated against the US real aggregate wealth data. Whereas the standard models explain the evolution of the aggregate wealth in the recent decades as an outcome of "rational bubbles" in which the model agents have good understanding of the structure of the economic system[9], the model here assumes that they know very little about it. It will be shown that the model can explain both qualitative and quantitative features of the actual data quite well. These include the durations and the amplitudes of contractions and expansions in the data. It will also be shown that the calibrated model exhibits a substantial departure from the outcome under Rational Expectations.

This paper complements the standard approaches for dealing with model uncertainty such as robust control theory and econometric learning theory. In robust control theory, a decision-maker deals with model uncertainty by deriving a decision rule that performs well across different model specifications. However, this is not possible here because the level of uncertainty is such that the model agents are not provided with a set of models to work with. The most comprehensive surveys of robust control theory are given by Hansen and Sargent (2008; 2010). For a shorter survey, see Williams (2008). For a historical viewpoint, see Stigler (2010). Some (non-exhaustive) recent

---

[8] See Barro and Sala-i-Martin (2003) and Acemoglu (2008) for textbook treatments of growth models.

[9] See Calvalho, Martin, and Ventura (2012) for an example.

works in this literature are Hansen, Sargent, Turmuhambetova, and Williams (2006), Cogley, Colacito, Hansen, and Sargent (2008), Hansen and Sargent (2005; 2007; 2011), and Ellison and Sargent (2012a; 2012b). In econometric learning theory, a decision-maker chooses a statistical model for objects of interest, estimates the parameters of this model using the actual data, and derives a decision rule based on the estimated model. This is not possible here because the information structure is such that the relevant data are not available to the model agents. The most comprehensive surveys of econometric learning theory are provided by Evans and Honkapohja (1999; 2001; 2009; 2011). Some (non-exhaustive) recent works in this literature are Sargent (1999), Cho, Sargent, and Williams (2002), Cogley and Sargent (2005; 2008), Sargent, Williams, and Zha (2006; 2009), Carceles-Poveda and Giannitsarou (2007), Ellison and Yates (2007), Cho and Sargent (2008), Williams (2009), Branch and Evans (2010; 2011), and Ellison and Pearlman (2011).

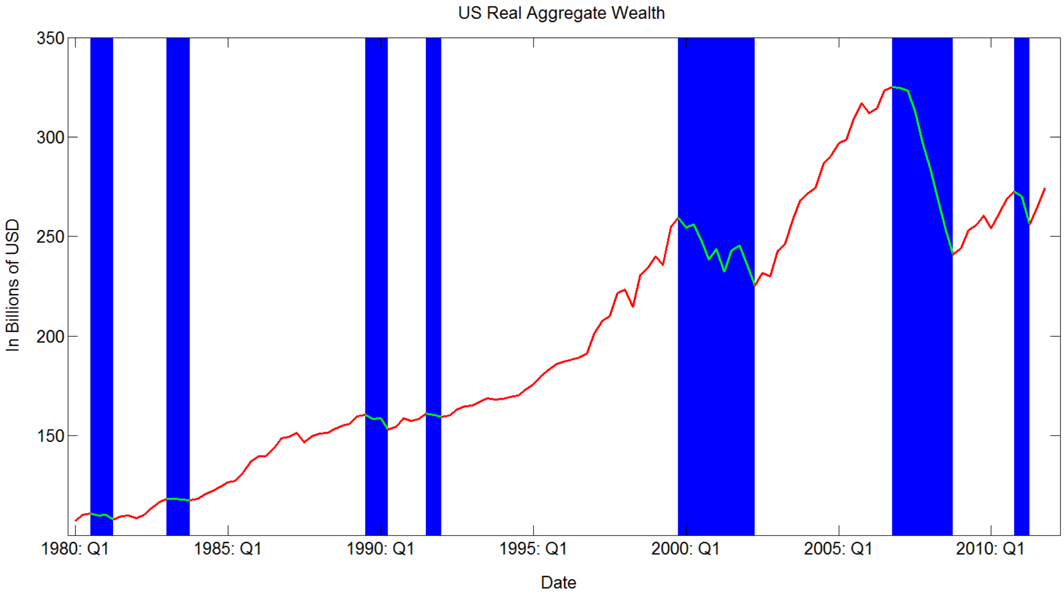
The model in this paper lies somewhere between the standard macroeconomic models (e.g., Dynamic Stochastic General Equilibrium models) and the agent-based computational models[10]. Whereas the models that belong to the former class derive the relevant equations from first principles (i.e., explicit optimization), the models that belong to the latter class typically impose the relevant equations as behavioral rules. This difference often implies a trade-off between rigor and flexibility. However, this is less of a problem with $ITEL$: even though it is a set of behavioral rules, following it is good in the sense that a system that is governed by it stays in PNEs of the relevant game a very high proportion of the time. Thus, the growth model in this paper consists of a set of behavioral rules that are quite optimal ($ITEL$) and a set of policy functions and transition equations that are derived from first principles.

The paper is structured as follows: section 3.2 studies the US real aggregate wealth data; section 3.3 presents the model; section 3.4 deals with calibration and simulation; section 3.5 concludes.

---

[10] See Gaffard and Napoletano (2012) for an overview of recent contributions to this literature. See Lengnick (2011), Dawid, Gemkow, Harting, van der Hoog, and Neugart (2012), and Fagiolo and Roventini (2012) for recent surveys of agent-based computational macroeconomic models.

Figure 3.1: US Real Aggregate Wealth (1980:Q1 to 2011:Q4)



The figure plots the US real aggregate wealth data from the first quarter of 1980 to the fourth quarter of 2011. The shaded areas correspond to contractions dated by the BBQ algorithm of Harding and Pagan (2002) (discussed in section 3.2).

3.2 US Real Aggregate Wealth Data

As mentioned above, the growth model in this paper will be evaluated by its ability to match the features of the US aggregate wealth data. Therefore, this section starts by taking a look at the data.

Figure 3.1 plots the US real aggregate wealth[11] from 1980:Q1 to 2011:Q4. One can see that the aggregate wealth went through tremendous changes in the recent decades, increasing by three-fold relative to the beginning of the 80's before the economic crisis of 2008 and falling abruptly thereafter. To add rigor to this casual visual analysis, the data are coursed through Harding and Pagan (2002)'s BBQ algorithm. The algorithm, which is an extension of Bry and Boschan (1971)'s business cycle dating algorithm for monthly data to quarterly data (thus the name BBQ), generates the business cycle dates in a time series by identifying its turning points (i.e., local troughs and peaks). Given that this paper aims to model both trend and cycles of aggregate

---

[11] It is obtained by deflating total net worth of households and non-profit organizations with consumer price index. The data are downloaded from Federal Reserve Bank of St Louis website (http://research.stlouisfed.org/).

variables, it is essential to work with a business cycle dating method which can handle these simultaneously. In contrast to more widely-used business cycle dating methods which rely on detrending of data (e.g., Hodrick and Prescott (1997)'s filter), the BBQ algorithm produces the business cycle dates using both trend and cycle components of data, i.e., the classical cycles which are cycles in the level of a variable. This is the reason why the BBQ algorithm is chosen over other business cycle dating methods here. There are other advantages as well. For instance, it is much easier to communicate the business cycles produced by the BBQ algorithm to the public as non-economists are almost always concerned with the classical cycles. For an exhaustive list of advantages, see Harding and Pagan (2002).

The BBQ algorithm produces the business cycle dates based on a set of censoring rules. The key rules are that (a) peaks and troughs alternate, (b) a phase (peak to the following trough or trough to the following peak) must last at least 2 quarters, and (c) a complete cycle (peak to the next peak or trough to the next trough) must be at least 5 quarters long. As explained by Harding and Pagan (2002), these rules can be thought of as a translation of visual judgment about the business cycles, which is good at identifying false turning points, into an algorithm in order to carry out the visual analysis more systematically.

The application of the BBQ algorithm to the US real aggregate wealth data is carried out using James Engel's code which implements the original BBQ algorithm more efficiently.[12] Table 3.1 gives the business cycle facts about the data based on the business cycle dates produced by the BBQ algorithm. "Expansion" refers to time periods between a trough and a peak that are side-by-side and "Contraction" time periods between a peak and a trough that are next to each other. It shows that on average (a) an expansion lasts longer than a contraction (row 1), (b) growth during an expansion is stronger than decline during a contraction (row 2), and (c) gains during an expansion outweigh losses during a contraction (rows 3 and 4). The shaded areas on Figure 3.1 correspond to the contractions dated by the BBQ algorithm. As foreshadowed by the

---

[12] It can be downloaded from http://www.ncer.edu.au/data/.

Table 3.1: Business Cycle Facts about US Real Aggregate Wealth (1980:Q1 to 2011:Q4)

Business Cycle Facts	Contraction	Expansion
Mean Duration (Quarters)	4.43	15.33
Mean Amplitude (%)	-8.48	23.86
Mean Cumulative Movement (%)	-27.14	253.21
Mean Cumulative Movement / Duration (%)	-3.96	12.36

The table summarizes the business cycle facts about the US real aggregate wealth data (shown in Figure 3.1) by using the BBQ algorithm of Harding and Pagan (2002). The algorithm produces business cycle dates in a time series by identifying its turning points (i.e. local troughs and peaks). Contraction refers to time periods between a peak and a trough that are side-by-side and Expansion time periods between a trough and a peak that are next to each other. Duration stands for the length of Contraction or Expansion, Amplitude the difference between the peak and the trough (decline) under Contraction or the difference between the trough and the peak (growth) under Expansion, and Cumulative Movement the sum of losses relative to the peak at the beginning of Contraction or the sum of gains relative to the trough at the beginning of Expansion. The table reports their time averages.

explanation above, these contractions accord well with what eyes pick out, especially those which are prolonged and large in magnitude (e.g., the contraction associated with the economics crisis of 2008).

The aim of the following section is to construct a model that can match the business cycle facts about the US real aggregate wealth data above.

3.3 Model

3.3.1 Setup

The economic system consists of n islands where $n \geq 2$. Each island is occupied by a continuum of infinitely-living households of unit measure. There is an interaction among the islands in the sense that the return on investment on each island depends on investment decisions made by the households on all the islands. However, the modes of interaction are not known to these households because the households on each island know nothing about other islands, including how many other islands are out there and what constraints the households on other islands face. Specifically, it is assumed that the mechanism which generates the return on investment is unknown to the households on all the islands, i.e., the return on investment is a black box[13]. This is how the Knightian Uncertainty (Knight, 1921), which refers to a situation in which the structure of a system is unknown to its agents, enters the economic system here. Because the assumption that the households on each island know nothing about other islands rules out learning, the mechanism is unknowable as well. Whereas the modes of inter-island interactions are unknown at all times, the modes of intra-island interactions are known at all times because every piece of intra-island information is perfectly observed by the households on the same island.

Every household has a Constant Relative Risk Aversion instantaneous utility function $u(c) = \frac{c^{1-\rho}}{1-\rho}$ where $c \geq 0$ is a consumption and ρ is the coefficient of relative risk aversion. The restriction that $0 \leq \rho < 1$ is imposed so that $u(0) = 0$.[14,15] It will be shown that this restriction makes the household's optimization problem more tractable (refer to subsubsection 3.3.3.1).

The households on each island are divided into f depositors and $1-f$ bankers where $0 < f < 1$. The depositors on each island allocate their wealth between consumptions and savings in each time period. Their savings are distributed equally among the bankers on the same island which invest them today for returns tomorrow on their behalf in each time period. It is assumed that the depositors cannot make intertemporal transfers of their wealth without going through the bankers because they do not have access to a reliable storage technology.

In each time period, there is a probability of $1 - \sigma$ that a depositor turns into a banker where $0 < \sigma < 1$. The population sizes of the depositors and the bankers are kept constant by assuming that $(1 - \sigma)f$ depositors and bankers switch their places in each time period. The restriction that $1 - f \geq (1 - \sigma)f$ is imposed so that there are enough banking jobs (whose number is $1 - f$ in total) to go around for depositors turning into bankers (whose number is $(1-\sigma)f$). The idea of modeling the economy as a set of islands and dividing households in depositors and bankers (and the associated

---

[13] This analogy is taken from Nax, Burton-Chellew, West, and Young (2013).

[14] Otherwise, $u(0)$ is not defined.

[15] The restriction that $0 \leq \rho \leq 1$ is for simplicity. The results in this paper can go through without this restriction (under some assumptions).

details) is taken from Gertler and Kiyotaki (2010). However, the model in this paper is very different from their model otherwise.

When a depositor turns into a banker, it (a) loses its wealth to a banker which takes its place as a new depositor and (b) receives nothing until it becomes a depositor again because the returns on investments are entirely captured by the depositors. Because $u(0) = 0$ is the minimum value of $u(c)$, it is clearly undesirable for a household to be a banker. However, the bankers still have an incentive to maximize the returns on investments. In each time period, a banker faces a probability of $\frac{(1-\sigma)f}{1-f}$ that it will become a depositor in the following time period (recall that $(1-\sigma)f$ is the number of bankers turning into depositors and $1-f$ is the total number of bankers). Because a depositor consumes a constant fraction of its wealth, which is positively related to the return on investment (refer to subsubsection 3.3.3.1), a banker's *ex-ante* utility is increasing in the return on investment. So, there is no conflict of interest between the depositors and the bankers.

In what follows, the depositors and the bankers on each island are modeled with the representative depositor and the representative banker respectively. The representative depositor is referred to as the Island Depositor (ID) and the representative banker as the Island Banker (IB) henceforth.

3.3.2 Island Banker's Problem

In each time period, the representative banker on each island, which is referred to as Island Banker (IB), makes an investment whose return is realized in the following time period. Although it has an incentive to maximize the return on investment (refer to subsection 3.3.1), the mechanism that generates the return is unknown and unknowable. The following subsubsections present a set of learning rules that is compatible with this extreme uncertainty and discuss its properties.

3.3.2.1 Investment Game and Knightian Uncertainty

This subsubsection introduces an investment game which generates the return on investment. Its main features are externalities and threshold effects which are discussed in the context of financial markets below.

There are two types of assets available for investment: the independent (I) and the dependent (D). IB on each island invests on behalf of the representative depositor on the same island, which is referred to as the Island Depositor (ID), by choosing between two portfolios[16]: L for a portfolio in which the proportion of D, which is denoted by α^L, is low and H for a portfolio in which the proportion of D, which is denoted by α^H, is high where $0 < \alpha^L < \alpha^H < 1$.[17] Whereas I has a constant gross rate of return, D has a gross rate of return which depends on how IBs on all the islands invest in it. These assumptions are the rationale behind the naming of the assets. Recall that n denotes the number of the islands. Because the bankers on each island are modeled by the representative banker IB, n is also the number of IBs. Let R^I_{t+1} be the gross rate of return on I in time period $t+1$, R^D_{t+1} the gross rate of return on D in time period $t+1$, r the maximum net rate of return on D ($\max R^D_{t+1} = 1 + r$ for all t) where $0 < r < 1$, and x_t the number of IBs which choose L in time period t. $R^I_{t+1} = R^I$ for all t and R^I is normalized to 1 for simplicity. R^D_{t+1} is generated by a mechanism

$$R^D(x_t) = \begin{cases} 1 + \left(\frac{n - x_t}{n}\right) r & \text{for } x_t \in \{0, \dots, \tau - 1\} \\ 1 + \left(\frac{x_t - n}{n}\right) r & \text{for } x_t \in \{\tau, \dots, n\} \end{cases} \quad \text{for all } t \tag{3.1}$$

where $\tau \in \{1, \dots, n-1\}$ is the threshold at which the coefficient of r changes its sign. The two main features of (3.1) are (a) externalities or spillovers and (b) threshold effects. These features have been considered to be important for modeling economic growth by various researchers (see Azariadis and Drazen, 1990; Azariadis and Stachurski, 2003) and this motivates their incorporation into the model in this paper.

---

[16] The restriction to two portfolio types is for simplicity. The model can be extended to any discrete number of portfolio types in a straightforward manner. In the next chapter, there will be ten portfolio types.

[17] This implies that IBs cannot divest completely from both I and D.

The dependence of R^D on x_t in (3.1) implies that the portfolio choice of an individual IB in time period t influences the gross rate of return on D all the IBs receive in time period $t+1$. The nature of this externality depends on where x_t is relative to the threshold τ. When the number of IBs choosing L is less than τ, more IBs choosing L decreases the gross rate of return on D (negative externality). When the number of IBs choosing L is greater than or equal to τ, more IBs choosing L increases the gross rate of return on D (positive externality). The functional form of R^D in (3.1) implies that the gross rate of return on D is greater than the gross rate of return on I when the number of IBs choosing L is less than τ and the other way around when it is greater than or equal to τ. This means that it is in IBs' interest to coordinate their portfolio choices to H (not choosing L altogether) if possible. When all the IBs choose L, D and I have the same gross rate of return of 1. It is easier to relate to R^I and $R^D(.)$ if one thinks of I as government debts and D as private debts and equities. In this case, the non-linear relationship between x and $R^D(x)$ in (3.1) can be considered as a model of the threshold effects in the financial markets[18,19]. The nature of the threshold effect above reflects the importance of market thickness[20] which is manifested in $R^D(.)$ being greater than R^I only when sufficiently many IBs choose to be highly involved with D. As shown by Hong, Kubik, and Stein (2004) and Brown, Ivkovik, Smith, and Weisbenner (2008), the stock market participation decision of an individual depends positively on her/his peers' participation rate. This suggests that an individual's stock market participation decision has externalities. Guiso, Sapienza, and Zingales (2008) provide evidence that there is a negative relationship between the lack of trust and the stock market participation. Moreover, they show that less trusting individuals buy less stock conditional on buying. Putting these together, the relationship between

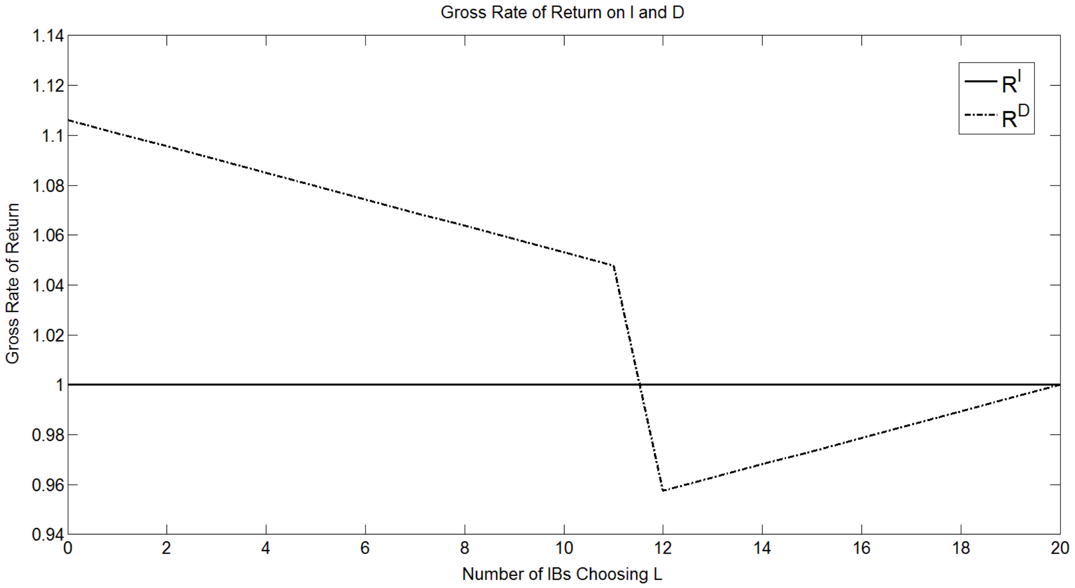
---

[18] Franses and van Dijk (2000) estimate various types of the threshold models using financial data such as stock returns and exchange rates and show that these are potentially useful for modeling and forecasting returns on financial assets. Hansen (2011) provides a brief literature review about the application of the threshold models to financial data such as interest rates, stock returns, and exchange rates.

[19] The next chapter provides empirical evidence of the threshold effects in the excess return of stocks over government bonds.

[20] McLaren (2003) and Roth (2009) discuss the importance of market thickness using a variety of examples. Stiglitz (1994) argues that government interventions in financial markets can be justified on the ground that thicker financial markets work better.

Figure 3.2: Gross Rate of Return on I and D

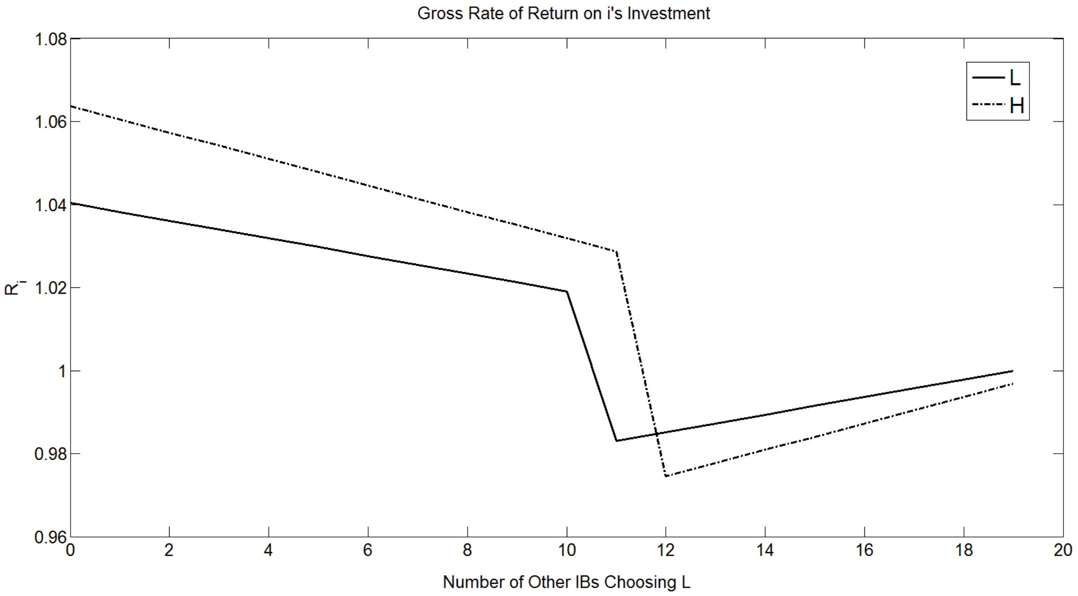


The figure plots the gross rate of return on Independent (R^I) and Dependent (R^D) against the number of IBs choosing portfolio L. R^I, which is set to 1, is plotted with the unbroken line and R^D, which follows (3.1), is plotted with the broken line. $n = 20$, $r = .1062$, and $\tau = 12$ which are the parameter values used in this paper.

the number of IBs choosing L and $R^D(.)$ above can be thought of as capturing the idea that the financial markets function better when more people choose to be highly involved in them because this signals that they are more trustworthy. Formally, (3.1) implies that (i) $R^D_{t+1} = R^D(x_t) > R^I_{t+1}$ and $R^D(x_t)$ is strictly decreasing in x_t when $x_t < \tau$ and (ii) $R^D_{t+1} = R^D(x_t) \leq R^I_{t+1}$ and $R^D(x_t)$ is strictly increasing in x_t when $x_t \geq \tau$ with $R^D_{t+1} = R^I_{t+1} = 1$ when $x_t = n$. Figure 3.2 plots both R^I and $R^D(.)$ against the number of IBs choosing L in order to illustrate (i) and (ii). R^I is plotted with the unbroken line and $R^D(.)$ with the broken line.

The IBs' portfolio choice problem in each time period is a game in the sense that each IB's gross rate of return on investment depends not only on its own portfolio choice but also on other IBs' portfolio choices. To see this more clearly, let a_{it} be the portfolio type that IB on island i, $i \in \{1, 2, \ldots, n\}$, chooses for investment in time period t. As mentioned above, it is assumed that the IB can choose from two portfolio types L and H so $a_{it} = L$ or H. Let $R_{i,t+1}$ be the associated gross rate of return on investment in the following time period. $R_{i,t+1} = \alpha^{a_{it}} R^D_{t+1} + (1 - \alpha^{a_{it}}) R^I_{t+1} =$

Figure 3.3: Gross Rate of Return on Investment for *IB* on Island *i*



The figure plots the gross rate of return on investment for *IB* on island i (R_i) against the number of other *IB*s choosing portfolio L. The unbroken line corresponds to the *IB* choosing L and the broken line to the *IB* choosing H. $n = 20$, $r = .1062$, $\tau = 12$, $\alpha^L = .4$, and $\alpha^H = .6$ which are the parameter values used in this paper.

$\alpha^{a_{it}} R^D_{t+1} + (1 - \alpha^{a_{it}}) = 1 + \alpha^{a_{it}}(R^D_{t+1} - 1)$ which is due to the normalization of R^I_{t+1} to 1. This brings out the interdependent nature of $R_{i,t+1}$ which arises from the dependence of R^D_{t+1} on x_t (the number of *IB*s choosing L). Figure 3.3 plots $R_{i,t+1}$ against the number of other *IB*s choosing L. The unbroken line corresponds to $a_{it} = L$ and the broken line $a_{it} = H$. As expected from its functional form, $R_{i,t+1}$ inherits the kinkedness of $R^D(.)$ which was shown in Figure 3.2. The static game which corresponds to the *IB*s' period-by-period portfolio choice problem is referred to as Investment Game (*IG*) henceforth. To be consistent with the game-theoretic terminologies, a_{it}s are called actions and $R_{i,t+1}$s payoffs from here onward.

Under common knowledge of rationality and common knowledge of the structure of the economic system, *IG* can be solved in the usual manner that a static game is solved. Let "coordinated outcomes" refer to situations in which every player in a game chooses the same action. In the context of *IG*, these correspond to $x_t = n$ (every *IB* chooses L) and $x_t = 0$ (every *IB* chooses H). Because *IG* is an n-person generalization

Table 3.2: 2-person Coordination Game

	L	H
L	$1, 1$	$1-\alpha^L\frac{r}{2}, 1-\alpha^H\frac{r}{2}$
H	$1-\alpha^H\frac{r}{2}, 1-\alpha^L\frac{r}{2}$	$1+\alpha^H r, 1+\alpha^H r$

This 2-person coordination game is produced by setting $n = 2$ and $\tau = 1$ in (3.1). Because $0 < \alpha^L < \alpha^H$ and $r > 0$, this game has multiple Pure Nash Equilibria which are (L, L) and (H, H).

of two-person coordination game[21] in Table 3.2, it is instructive to analyze the latter first. The 2-person coordination game is produced by setting $n = 2$ and $\tau = 1$ in (3.1). Because $0 < \alpha^L < \alpha^H$ and $r > 0$, this game has multiple Pure Nash Equilibria (PNEs) which are (L, L) and (H, H). As defined above, these are the coordinated outcomes in this game with (L, L) corresponding to $x_t = 2$ and (H, H) to $x_t = 0$. This result carries over to IG: it can be shown that the coordinated outcomes, $x_t = n$ and $x_t = 0$, are the only PNEs[22] in IG (refer to Appendix 3.6.1). Figure 3.3 provides some intuitions for this result. It shows that when all other IBs choose H (the beginning of the horizontal axis), choosing H (the broken line) gives a higher payoff than choosing L (the unbroken line). It also shows that when all other IBs choose L (the end of the horizontal axis), choosing L (the unbroken line) yields a higher payoff than choosing H (the broken line). Thus, $x_t = 0$ and $x_t = n$ are the PNEs of IG. The figure also brings out that the threshold effect, which is associated with the change of sign of the coefficient of r in (3.1) the neighborhood of τ, is the reason for the existence of multiple PNEs in IG. Note that conditional on a particular PNE being reached, this PNE is also a Rational Expectations Equilibrium (REE). To see this concretely, suppose that $x_t = n$. This PNE is also an REE in the sense that the subjective probability distribution (i.e., every IB believes that all other IBs choose H so it chooses H) and the equilibrium probability distribution (i.e., every IB chooses H) coincide.[23]

---

[21] See Cooper (1998) for a textbook treatment of coordination games in macroeconomics.

[22] A Mixed Nash Equilibrium is not considered in this paper.

[23] See Sargent (2008) for this definition of Rational Expectations Equilibrium.

Now, consider an infinitely repeated game in which IG is a stage game. The outcomes under which IBs are in a PNE in all stage games (i.e. $x_t = 0$ or $x_t = n$ for all t) can be supported as Subgame Perfect Nash Equilibria ($SPNE$s).[24] Intuitively, this is because the deviation by any IB in a stage game leads to its own payoff loss. Conditional on a particular $SPNE$ being reached, this $SPNE$ is also a time-consistent REE. The time-consistency aspect is due to the notion of Subgame Perfection which guarantees that following this $SPNE$ strategy is optimal at any point in the repeated game. The REE aspect is for the same reason as in the previous paragraph.

However, it is assumed that the specification of $R^D(.)$ is unknown to all the households on all the islands for all t. This implies that the common knowledge breaks down as the households performing the role of IB do not know how their actions map into their payoffs. Because the households on each island know nothing about other islands, $R^D(.)$ is unknowable as well as learning about it requires information about other islands. In this case, the PNEs in IG cease to be a useful solution concept. The same is true for the $SPNE$s of the infinitely repeated game in which IG is the stage game.

Knightian Uncertainty (Knight, 1921) can be interpreted as a situation in which the structure of a system is unknown to its agents. Because such is the case for the households on all the islands, it can be said that they face Knightian Uncertainty.

3.3.2.2 Interactive Trial and Error Learning

Given that IBs do not know how their actions determine their payoffs, how should they decide on their actions? As mentioned in subsubsection 3.3.2.1 above, they cannot learn this process as well because their information sets contain only their own actions and payoffs. Here, it is assumed that they carry on under this uncertainty by following a set of adaptive learning rules called Interactive Trial and Error Learning ($ITEL$)

---

[24] Friedman (1971)'s Folk Theorem suggests that other outcomes can be supported as Subgame Perfect Nash Equilibria as well. But, these are asymmetric. Given that macroeconomists are usually only concerned with symmetric equilibria (mainly due to the representative agent assumption), this paper only considers these as equilibria. Some game theorists argue that symmetric equilibria are more compelling than asymmetric equilibria in a symmetric game like IG here (see Dutta, 1999).

(Young, 2009). $ITEL$ is suitable for this environment because it depends only on one's own actions and payoffs.[25]

$ITEL$ has two search phases: (a) active search in which an agent experiments with an action that is different from its current benchmark action and accepts it if it results to a payoff higher than its current benchmark payoff and (b) passive search in which an agent experiments with an action, not necessarily different from its current benchmark action, and accepts it with a probability that is strictly increasing in the payoff from that action and strictly decreasing in its current benchmark payoff. These two also differ in how they get activated: while the former is self-induced, the latter is induced by other agents.

$ITEL$ has a number of nice properties. First, it is fairly intuitive. Second, it is not based on an explicit model which makes it broadly applicable. Third, as the probability of the active search, which is denoted by ε, becomes arbitrarily small (but still positive), $ITEL$ selects PNEs in a stage game of an infinitely repeated game a very high proportion of the time.[26] Whereas the first two properties will become obvious once $ITEL$ is stated below, the third property deserves more discussion. It says that following $ITEL$ guarantees an outcome of the stage game under rationality (i.e., common knowledge of game), which is a PNE, to prevail very frequently. As discussed in subsubsection 3.3.2.1, a PNE of the stage game also corresponds to a REE which is the benchmark equilibrium concept in macroeconomics. In this sense, the outcome under $ITEL$ can be thought of as a measurable deviation from the outcome under Rational Expectations with the extent of the deviation governed by ε. The relationship between ε and the frequency of PNE outcomes will be investigated using simulations in subsection 3.4.3. Note that the source of the deviation is Knightian Uncertainty: this is the reason why $ITEL$ is being followed in the first place.

Even though agents are in PNEs of a stage game most of the time under $ITEL$

---

[25] Adaptive learning rules which depend only on one's own actions and payoffs are called completely uncoupled learning rules. See Young (2009) for a brief literature review.

[26] This is equivalent to a statement that PNEs are Stochastically Stable States. See Young (2009) for a precise definition of this statement and assumptions required for this result to follow. Pradelski and Young (2012) extend this result so that only efficient PNEs are Stochastically Stable States.

(when ε is very small), they occasionally move out of them as a consequence of learning. This is similar to Escape Dynamics which occasionally moves a model out of equilibrium because agents learn based on mis-specified models[27]. However, there is a crucial difference between the two: while the deviation under $ITEL$ is concerned with a Nash Equilibrium, Escape Dynamics is concerned with a Self-confirming Equilibrium (Fudenberg and Levine, 1993) which may not coincide with the Nash Equilibrium[28]. The deviations differ in their origins as well: while the deviation under $ITEL$ is a consequence of deliberate experimentations by agents (see below), Escape Dynamics arises due to changes of beliefs by agents.

In what follows, $ITEL$ is reproduced for the ease of reference. Young (2009)'s notations are used as much as possible to preserve his originality. Time subscripts t and $t-1$ are used instead of $t+1$ and t to emphasize the backward-looking nature of $ITEL$.

Let A_i stand for the set of possible actions of IB on island i (i.e., $A_i = \{L, H\}$ for all i). $z_{i,t-1} = (m_{i,t-1}, \bar{a}_{i,t-1}, \bar{R}_{i,t-1})$ is a state vector of IB on island i in time period $t-1$ where $m_{i,t-1}$ is its mood, $\bar{a}_{i,t-1}$ its benchmark action, and $\bar{R}_{i,t-1}$ its benchmark payoff. There are four moods: content (c), discontent (d), hopeful (h), and watchful (w).[29] The learning rules are as follows:[30]

- $m_{i,t-1} = c$: In $t-1$, IB on island i experiments with an action that is different from the benchmark action with a probability of ε (active search). If it experiments, a new action is drawn from $A_i - \{\bar{a}_{i,t-1}\}$ uniformly at random.
  - $a_{i,t-1} \neq \bar{a}_{i,t-1}$ and $R_{it} \leq \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.
  - $a_{i,t-1} \neq \bar{a}_{i,t-1}$ and $R_{it} > \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = a_{i,t-1}$, and $\bar{R}_{it} = R_{it}$.
  - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $R_{it} = \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.

---

[27] Escape Dynamics has been used to explain inflation dynamics in various contexts. See Sargent (1999), Cho, Williams, and Sargent (2002), Ellison and Yates (2007), and Sargent, Williams, and Zha (2009) for examples. Ellison and Scott (2013) provide a novel application of this mechanism to analyzing non-renewable resource markets. See Williams (2002) and Williams (2009) for a formal analysis of Escape Dynamics.

[28] See Fudenberg and Levine (1993) and Williams (2002) for this point.

[29] Young (2009) interprets the mood as a psychological state.

[30] See Young (2009) for graphical representations of $ITEL$.

  - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $R_{it} > \bar{R}_{i,t-1} \Rightarrow m_{it} = h$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.
  - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $R_{it} < \bar{R}_{i,t-1} \Rightarrow m_{it} = w$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.

- $m_{i,t-1} = h$: In $t-1$, *IB* on island i maintains its benchmark action.
  - $R_{it} = \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.
  - $R_{it} > \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = R_{it}$.
  - $R_{it} < \bar{R}_{i,t-1} \Rightarrow m_{it} = w$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.

- $m_{i,t-1} = w$: In $t-1$, *IB* on island i maintains its benchmark action.
  - $R_{it} = \bar{R}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.
  - $R_{it} > \bar{R}_{i,t-1} \Rightarrow m_{it} = h$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.
  - $R_{it} < \bar{R}_{i,t-1} \Rightarrow m_{it} = d$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$.

- $m_{i,t-1} = d$: In $t-1$, *IB* on island i draws $a_{i,t-1}$ from A_i uniformly at random (passive search).
  - $m_{it} = c$, $\bar{a}_{it} = a_{i,t-1}$, and $\bar{R}_{it} = R_{it}$ with a probability of $\phi(R_{it}, \bar{R}_{i,t-1})$.
  - $m_{it} = d$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{R}_{it} = \bar{R}_{i,t-1}$ with a probability of $1-\phi(R_{it}, \bar{R}_{i,t-1})$.

As mentioned above, $\phi(R_{it}, \bar{R}_{i,t-1})$ is strictly increasing in R_{it} and strictly decreasing in $\bar{R}_{i,t-1}$. This monotonicity requirement is very intuitive: it implies that an *IB* is more likely to be content with its action if it delivers a higher payoff but is less likely to be content with its action if its reservation payoff is higher. Young (2009) shows that the precise functional form of $\phi(.)$ is not important for *ITEL*'s ability to select *PNE*s a high proportion of the time. Here, the functional form

$$\phi(R_{it}, \bar{R}_{i,t-1}) = e^{\lambda((R_{it} - \bar{R}_{i,t-1}) - \omega)} \tag{3.2}$$

is used where $\lambda > 0$ and ω is such that $0 < \phi(.) < 1$. Setting ω to a value greater than the maximum payoff difference, which is $\alpha^H(2n - \tau)\frac{r}{n}$, ensures that the latter restriction holds.

For some, a more intuitive understanding of $ITEL$ may be achieved by approaching it as a numerical optimization program that is designed to explore the domain of an objective function more thoroughly. The usual problem with an optimization program is that it can get stuck at a local optimum.[31] The active search and the passive search of $ITEL$ guarantee that this does not happen with them as other points far away from the current position in the domain will be explored eventually. This is similar in spirit to Simulated Annealing (Kirkpatrick, Gelatt, and Vecchi, 1983).[32] Of course, $ITEL$ is not an optimization program in the conventional sense because the presence of the active search implies that $ITEL$ does not have termination criteria.

Before proceeding further, it is worth re-emphasizing that $ITEL$ is used as a behavioral rule here.[33]

3.3.3 Island Depositor's Problem

In each time period, the representative depositor on each island, which is referred to as the Island Depositor (ID), allocates its wealth between consumption and saving. Because of Knightian Uncertainty[34], it faces a difficulty with forming expectations about the future which are necessary for the wealth allocation in the current time period. Specifically, Knightian Uncertainty implies that it is not obvious to IDs what their conditional expectations in t, denoted by $\tilde{E}_t$, should be. The following subsubsections introduce features that allow IDs to set up and solve their optimal consumption/saving problem and discuss the resulting dynamics.

---

[31] This happens frequently in the estimation of macroeconomic models. For instance, Fernandez-Villaverde and Rubio-Ramirez (2007) discuss that in estimating their model, they fix the depreciation factor because the likelihood function is flat in this dimension (i.e., the depreciation factor is not well-identified).

[32] See Rutenbar (1989) and Bertsimas and Tsitsiklis (1993) for an overview.

[33] Behavioral rules are not too uncommon in macroeconomics. Taylor Rule (Taylor, 1993) is one prominent example.

[34] Recall that there is a probability of $1-\sigma$ that a depositor turns into a banker in each time period. Assuming an arbitrarily long history, a depositor should have been a banker at some point in the past. So, the depositor should be aware of Knightian Uncertainty associated with Investment Game (IG) in subsubsection 3.3.2.1.

3.3.3.1 Optimal Policy Functions and Transition Equation

In order to formulate ID's problem, it is necessary to define a timeline of events first:

1. At the beginning of the current time period, ID and IB on each island observe what payoff is realized from IB's action in the previous time period.
2. IB transfers the entire return (the payoff in the current time period × ID's saving in the previous time period) to ID and this becomes ID's wealth in the current time period.
3. $(1-\sigma)f$ depositors and bankers switch their places.
4. IB chooses what action to take.
5. ID observes what action is chosen by IB and allocates its wealth between consumption and saving. The saving is transferred to IB.
6. IB carries out its action.

This timeline of events repeats itself in each time period. It is summarized in Figure 3.4.

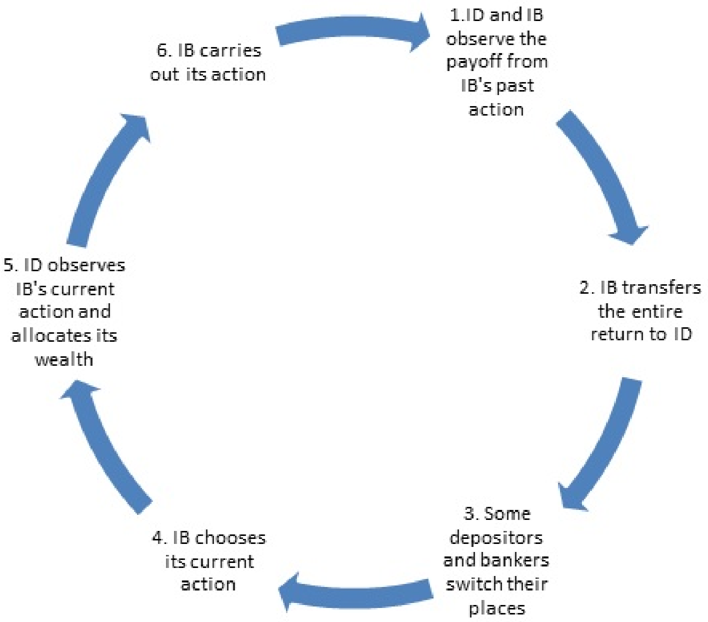
Recall that there is perfect information within an island. When ID allocates its wealth between consumption and saving (step 5 above), its information set contains a history of IB's (a) state, (b) action, and (c) payoff up to the current time period. This information set contains nothing about other islands.

Given this information set, ID on island i in time period t solves the following optimal consumption/saving problem:

$$\max_{\{c_{is}\}_{s=t}^{\infty}} \tilde{E}_t[\sum_{s=t}^{\infty}(\sigma\beta)^{s-t}\frac{c_{is}^{1-\rho}}{1-\rho}]\ s.t.\ w_{i,s+1} = R_{i,s+1}(w_{i,s}-c_{i,s}) \text{ for } s \geq t \text{ and a given value of } w_{it},$$

where $c_{i,s}$ is its consumption in s, $\tilde{E}_t$ its expectation conditional on the information set in t, $0 < \beta < 1$ its discount factor, w_{is} its wealth in s, and $R_{i,s}$ its payoff in s. Recall that the restriction of $0 \leq \rho < 1$ was imposed in subsection 3.3.1 so that $u(0) = 0$. This makes the mathematical representation of ID's optimal control problem simpler

Figure 3.4: Timeline of Events



The figure summarizes the timeline of events involving ID and IB on the same island. The number indices on the figure correspond to those in the text.

as one does not need to keep track of the utility associated with turning into a banker: it is zero because a banker receives the zero level of consumption and $u(0) = 0$.

As it was discussed above, it is not so obvious what $\tilde{E}_t$ should be under Knightian Uncertainty. Here, three beliefs are imposed on $\tilde{E}_t$ in order to make it determinate: ID believes that (a) its information set in the current time period contains all possible payoffs, (b) IB's current action is going to be played forever, and (c) the worst payoff will result from any action. These three beliefs turn the optimal consumption/saving problem under uncertainty[35] into one under certainty which is easier to solve. (a) transforms the set of observed payoffs into a support of some probability distributions. (b) restricts the support that results from (a) to the payoffs that are associated with the current action. (c) further restricts the support that arises from (a) and (b) to a singleton. When ID forms its expectation with these three beliefs, it always ends up with the imagined certainty. This is similar to Anticipated Utility approach of Kreps (1998) in which an economic agent optimizes its objective function taking the latest estimates of the relevant parameters as correct and permanent and updates them as it

---

[35] See Miao (2004) for a thorough analysis of this problem in a similar context.

goes along.[36]

Unlike (a) and (b), (c) is not very essential for the main results of this section. This can be replaced by any other belief that selects a unique independent and identically distributed probability distribution. Also, (c) can be replaced by Gilboa-Schmeidler preference (Gilboa and Schmeidler, 1989) under which an agent selects a probability distribution that gives the minimum expected value of its objective function because it is ambiguity-averse.[37] This probability distribution is the one implied by (c).

For simplicity, it is assumed that ID's information set actually contains all possible payoffs. Given $a_{it} = L$ or H, ID's problem is

$$\max_{\{c_{is}\}_{s=t}^{\infty}} \sum_{s=t}^{\infty} (\sigma\beta)^{s-t} \frac{c_{is}^{1-\rho}}{1-\rho}\; s.t.\; w_{i,s+1} = R_{\min}^{a_{it}}(w_{i,s} - c_{i,s}) \,\text{for}\, s \geq t \,\text{and a given value of}\, w_{it}$$

where $R_{\min}^{a_{it}}$ is the minimum payoff associated with a_{it}. (3.1) implies that $0 < R_{\min}^{H} < R_{\min}^{L} < 1$.[38] The associated Bellman equation is

$$V^{a_{it}}(w_{it}) = \max_{c_{it} \in [0, w_{it}]} [\frac{c_{it}^{1-\rho}}{1-\rho} + \sigma\beta V^{a_{it}}(w_{i,t+1})]\; s.t.\; w_{i,t+1} = R_{\min}^{a_{it}}(w_{i,t} - c_{i,t}) \,\text{for a given}\, w_{it}$$

where $V^{a_{it}}(.)$ is the value function. It is a standard result that as long as $\sigma\beta(R_{\min}^{a_{it}})^{1-\rho} < 1$, $V^{a_{it}}$ is well-defined. Because $0 < \sigma, \beta, R_{\min}^{L}, R_{\min}^{H}, \rho < 1$, this is the case here.[39] The optimal policy functions and the wealth transition equation are as follows:[40]

$$c_{it} = (1 - (\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}})w_{it}; \tag{3.3}$$

$$s_{it} = (\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}} w_{it}; \tag{3.4}$$

---

[36] The Anticipated Utility is popular among researchers in Econometric Learning literature because it saves on computational sophistication. See Ellison and Yates (2007) for an example of this point.

[37] Gilboa-Schmeidler preference underlies Robust Control Theory which is the latest version of optimal control theory that takes into account model specification errors in formulating an optimal policy. As mentioned in section 3.1, see Hansen and Sargent (2010) for a nice introduction to the literature.

[38] Given the belief that the worst payoff will result from any action, ID would not entrust its saving with IB if it had access to a reliable storage technology. But, such is ruled out by assumption (refer to subsection 3.3.1).

[39] Note that it is sufficient to have $\sigma\beta(R_{\min}^{L})^{1-\rho} < 1$ because $\sigma\beta(R_{\min}^{H})^{1-\rho} < \sigma\beta(R_{\min}^{L})^{1-\rho}$.

[40] These are standard results which can be found in many optimization textbooks for economists (e.g., Dixit, 1990).

$$w_{i,t+1} = R_{i,t+1}(\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}} w_{it}. \tag{3.5}$$

(3.3) and (3.4) make clear that the optimal policy functions depend on the current action of IB. (3.3), which can be rewritten as $\frac{c_{it}}{w_{it}} = 1 - (\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}}$, says that ID consumes a larger proportion of its current wealth when IB's current action is H than when it is L because $0 < R_{\min}^{H} < R_{\min}^{L}$ (the downside associated with H is worse than the downside associated with L).

3.3.3.2 Wealth Dynamics

(3.5) can be rewritten as

$$\frac{w_{i,t+1}}{w_{it}} = R_{i,t+1}(\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}} \tag{3.6}$$

which is the gross rate of change for the wealth of ID on island i in time period $t+1$. Whether (3.6) is less than 1 or not depends on the relative magnitudes of $R_{i,t+1}$ and $(\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}}$ which are determined by $a_t = (a_{it}, \ldots, a_{nt})$. Because $(\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}} < 1$, the role of $R_{i,t+1}$ is more important for this. (3.1) implies that $R_{i,t+1} \leq 1$ for $x_t \geq \tau$, so (3.6) is less than 1 in these cases. $x_t = n$ is particularly interesting among these. Recall that $ITEL$ selects the PNEs of IG a high proportion of the time given an arbitrarily small positive value of ε. But, $x_t = n$ is a PNE of IG. So, it is possible that the wealth declines for a protracted period of time under $ITEL$. (3.1) also implies that $R_{i,t+1} > 1$ when $x_t < \tau$. But, in order for (3.6) to be greater than or equal to 1, x_t should be sufficiently small so that $R_{i,t+1}$ is sufficiently large. The parameter values will be such that (3.6) is greater than 1 at least when $x_t = 0$ which is the other PNE of IG that $ITEL$ selects a high proportion of the time. So, ID's wealth can go through protracted periods of decline and growth under $ITEL$. This in turn implies that the fluctuations in the growth rate of ID's wealth usually occur when IBs' actions are not aligned with the PNEs of IG, i.e., the system's state is consistent with the disequilibria of IG.

The aggregate wealth in time period $t+1$ is denoted by W_{t+1} and takes the form

$$W_{t+1} = f \sum_{i=1}^{n} w_{i,t+1}$$

because there are f depositors on each island. The gross rate of change for the aggregate wealth is

$$\frac{W_{t+1}}{W_t} = \frac{\sum_{i=1}^{n} w_{i,t+1}}{\sum_{i=1}^{n} w_{it}}$$

whose functional form suggests that it inherits the properties of (3.6) in the PNEs. Consequently, the aggregate wealth can go through protracted periods of decline and growth under $ITEL$ too. This also implies that the fluctuations in the growth rate of the aggregate wealth usually occur when the system's state is consistent with the disequilibria of IG.

3.3.3.3 Consumption Dynamics

Even though the consumption dynamics is not the focus of this section, it is presented for the sake of completeness. Based on (3.3), the gross rate of change for the consumption of ID on island i in time period $t+1$ is

$$\frac{c_{i,t+1}}{c_{it}} = \frac{(1-(\sigma\beta(R_{\min}^{a_{i,t+1}})^{1-\rho})^{\frac{1}{\rho}})}{(1-(\sigma\beta(R_{\min}^{a_{it}})^{1-\rho})^{\frac{1}{\rho}})} \frac{w_{i,t+1}}{w_{i,t}} = \psi(a_{it}, a_{i,t+1}) \frac{w_{i,t+1}}{w_{i,t}} \tag{3.7}$$

which depends not only on a_{it} but also on $a_{i,t+1}$. In what follows, all possibilities for a_{it} and $a_{i,t+1}$ are considered.

When $a_{i,t+1} = a_{it}$, (3.6) and (3.7) are equal because $\psi(a_{it}, a_{i,t+1}) = 1$. Because (3.6) is less than 1 when $x_t \geq \tau$, (3.7) is less than 1 for these cases. In particular, (3.7) is less than 1 when $x_t = n$ which is a PNE of IG. When $x_t = 0$, which is also a PNE of IG, (3.7) is greater than 1 because the system is parameterized in such a way that (3.6) is greater than 1 at least for this case. Because $ITEL$ selects the PNEs of IG a very high proportion of the time for an arbitrarily small positive value of ε, ID's consumption, just like its wealth, can go through long periods of decline and growth.

This also implies that the fluctuations in the growth rate of ID's consumption usually occur when the system's state is consistent with the disequilibria of IG.

There are two cases to consider when $a_{i,t+1} \neq a_{it}$. When $a_{it} = L$ and $a_{i,t+1} = H$, the effect of (3.6) on (3.7) is reinforced because $\psi(L, H) > 1$ whereas when $a_{it} = H$ and $a_{i,t+1} = L$, it is attenuated because $\psi(H, L) < 1$.

The aggregate consumption in time period $t + 1$ is C_{t+1} which has the form

$$C_{t+1} = f \sum_{i=1}^{n} c_{i,t+1}$$

because there are f depositors on each island. Its gross rate of change is

$$\frac{C_{t+1}}{C_t} = \frac{\sum_{i=1}^{n} c_{i,t+1}}{\sum_{i=1}^{n} c_{it}}$$

which also inherits the properties of (3.7) in the PNEs. Like the aggregate wealth, the aggregate consumption can go through long periods of decline and growth under $ITEL$. This also implies that the fluctuations in the growth rate of the aggregate consumption usually occur when the system's state is consistent with the disequilibria of IG.

3.3.4 Discussion

Because $ITEL$ selects the PNEs of IG a very high proportion of the time, the aggregate wealth and the aggregate consumption can experience protracted periods of decline ($x_t = n$) or growth ($x_t = 0$). Because the growth rate of the aggregate wealth and the aggregate consumption are equal and constant during these periods, the economy can be said to be on a temporary balanced-growth-path. However, these periods eventually come to end because $ITEL$ requires some agents to engage in the active (experimental) search (refer to subsubsection 3.3.2.2) for an action that can give a higher payoff. Upon exiting the balanced-growth periods, the growth rates of the aggregate variables continue to fluctuate until the system comes to a rest again. This temporary resting point is not necessarily consistent with the PNEs of IG but is likely to be so. So, the fluctuations in the growth rates of the aggregate variables are usually associated with the disequilibria of IG.

3.4 Calibration and Simulation

3.4.1 Calibration

The model in this paper is calibrated by the Simulated Method of Moments[41]. Let W_t^a denote the aggregate wealth in time period t from the actual data (US real aggregate wealth from 1980:Q1 to 2011:Q4) and $W_k^s(\theta)$ the aggregate wealth in time period k generated by the model in this paper based on a vector of parameters $\theta = (\theta_1, \theta_2)$ where $\theta_1 = (n, r, \tau, \varepsilon, \lambda)$ is a sub-vector of parameters that are calibrated to the actual data and $\theta_2 = (\rho, f, \sigma, \alpha^L, \alpha^H, R^I, \beta)$ is a sub-vector of parameters that are fixed at certain values. Let $m^{BBQ}(.)$ be a function that maps data (actual or model-generated) into a 6 by 1 column vector whose entries are mean duration, mean amplitude, and mean cumulative movement of contractions and expansions respectively and $s^{BBQ}(.)$ a function that maps data into a 6 by 1 column vector whose entries are the corresponding standard deviations. Both $m^{BBQ}(.)$ and $s^{BBQ}(.)$ are based on the BBQ algorithm (refer to section 3.2) as the superscript suggests. The calibrated value of θ_1 is

$$\hat{\theta}_1 = \arg\min_{\theta_1} \sum_{l=1}^{6} \left(\frac{m_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a}) - m_l^{BBQ}(\{W_k^s((\theta_1, \theta_2))\}_{k=1}^{T^s})}{s_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a})} \right)^2$$

for a fixed value of θ_2 where T^a is the length of the actual data and T^s the length of the model-generated data.[42] The objective function above implies that the entries of $m_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a})$ whose standard deviations are relatively higher have relative lower weights in it. The minimization above is identical to what takes place under the Simulated Method of Moments where the weighting matrix is a diagonal matrix whose diagonal entries are the elements of $s_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a})$. However, the calibration above is

---

[41] Lee and Ingram (1991) and Duffie and Singleton (1993) introduced the Simulated Method of Moments in the context of dynamic models. See Adda and Cooper (2003), Canova (2007) De Jong and Dave (2011) for a textbook treatment of simulation-based estimation methods at large.

[42] Adam, Marcet, and Nicolini (2007) calibrate their model in a similar manner.

Table 3.3: Parameter Values

Parameter	Value	Type
n	20	Calibrated
r	.1062	Calibrated
τ	12	Calibrated
ε	.004	Calibrated
λ	39.5075	Calibrated
ρ	.9999	Fixed
f	.9	Fixed
σ	.99	Fixed
α^L	.4	Fixed
α^H	.6	Fixed
R^I	1	Fixed
β	.9999	Fixed

The table provides the parameter values used in this section. Calibrated parameters correspond to θ_1 and Fixed parameters correspond to θ_2 in the text.

not estimation because the matching function involved is not differentiable with respect to the parameters involved. In the calibration, $T^a = 128$ and $T^s = 500 \times T^a = 64000$. Table 3.3 gives the values of the calibrated parameters, which are denoted by θ_1, and the values of the fixed parameters, which are denoted by θ_2. Refer to Appendix 3.6.2 for further details about the calibration.

3.4.2 Business Cycle Facts from Simulation

Table 3.4 presents the business cycle facts about the model-generated aggregate wealth data using the BBQ algorithm (refer to section 3.2 for the details). The model is parameterized at the values in Table 3.3. The reported numbers are averages of time averages across 1000 simulation rounds, with the length of each round set to the length of the actual US real aggregate wealth data (128 periods) after discarding first 1000 periods to reduce the influence of the initial point (refer to Appendix 3.6.2 for more about this). The business cycle facts about the actual data are given in Table 3.4 as well for the ease of reference.

Columns 2 and 3 give the business cycle facts about the model-generated data. These are qualitatively similar to the business cycle facts about the actual data which are given in columns 4 and 5. These are quantitatively similar as well even though

Table 3.4: Business Cycle Facts

Business Cycle Facts	Model Contraction	$T^s = 128$ Expansion	Actual Contraction	$T^a = 128$ Expansion
Mean Duration (Quarters)	3.87	16.17	4.43	15.33
Mean Amplitude (%)	-6.74	24.91	-8.48	23.86
Mean Cumulative Movement (%)	-18.53	364.27	-27.14	253.21

The table gives the business cycle facts about the model-generated aggregate wealth data which correspond to the entries under Model. These are averages of time averages across 1000 simulation rounds. The business cycle facts about the actual US real aggregate wealth data, which are the entries under Actual, are also reported for the ease of reference. These are time averages. T^s is the length of each simulation round and T^a is the length of the actual data.

the magnitudes of the business cycle measures are understated for contraction and overstated for expansion. Overall, the model in this section can successfully match several business cycle facts about the US real aggregate wealth data.

3.4.3 Discussion

Recall that IG is the stage game of the infinitely repeated game played by IBs. As mentioned in subsubsection 3.3.2.2, as ε becomes an arbitrarily small positive number, $ITEL$ implements the PNEs of IG a very high proportion of the time. However, Young (2009) does not provide much guide on how small ε should be for this result to hold. As an attempt to shed light on this issue, Table 3.5 shows how often the model in this section is in the PNEs of IG for various values of ε. Other parameters are set at the values in Table 3.3.[43] Everything else is the same as in subsection 3.4.2 (1000 simulation rounds, each round 128 periods long after discarding first 1000 periods, etc.).

As expected, there is a negative relationship between ε and the proportion of the time in the PNEs of IG. Recall that the $SPNE$s of the infinitely repeated game in which IG is the stage game involve being in a PNE of IG for all time periods. Moreover, these are consistent with time-consistent REEs (refer to subsubsection 3.3.2.1). As shown in Table 3.5, $\varepsilon = .004$ obtained from the calibration implies that the business

---

[43] Strictly speaking, this is not optimal because the optimal values of these parameter (in terms of calibration) should depend on what value ε takes.

Table 3.5: Proportion of Time in PNEs of IG

ε	Proportion
.001	.9061
.002	.9039
.003	.7978
.004	.6699
.005	.5815

Proportion stands for the proportion of the time the model in this section is in the PNEs of IG. ε is the probability of the active (experimental) search defined in subsubsection 3.3.2.2. The reported numbers are averages of 1000 simulation rounds.

cycles contained in the US real aggregate wealth data from 1980:Q1 to 2011:Q4 are consistent with substantial departures from the outcomes under Rational Expectations (around 30% of the time). Note that the calibration in subsection 3.4.1 is essentially matching the disequilibrium states of IG against the actual data because the business cycles are not defined when the model is resting at a particular point x_t (the number of IBs choosing portfolio L in time period t; refer to subsubsection 3.3.2.1) which is likely to be a PNE of IG.

3.5 Conclusion

This paper is concerned with the growth of the aggregate variables under Knightian Uncertainty. Because the model agents do not know the structure of the economic system under Knightian Uncertainty, they cannot derive the optimal policy function for choosing investment portfolios. Here, they follow Interactive Trial and Error Learning ($ITEL$) (Young, 2009) as behavioral rules for choosing their investment portfolios. The key feature of $ITEL$ is that players learn about their environments by experimenting with new actions. This learning-by-experimenting is the reason behind the fluctuations of the aggregate variables in this paper.

The key implication of the growth model in this paper is that even though the aggregate variables are on the balanced-growth-path most of the time, there are recurrent fluctuations in their growth rates which kick them off the balanced-growth-path. The model is calibrated against the US real aggregate wealth. It can successfully match several business cycle features of the actual data such as the durations and the amplitudes of the contractions and the expansions.

3.6 Appendix

3.6.1 Pure Nash Equilibria of Investment Game

Recall that Investment Game (IG) is a static game in which Island Bankers' (IBs) actions in time period t determine their payoffs in time period $t+1$. There are n IBs each of which invests on behalf of Island Depositors (IDs) on the same island and $n \geq 2$ by assumption. The action of IB on island i in time period t, which is denoted by a_{it}, takes on either L or H and $a_t = (a_{1t}, \ldots, a_{nt})$ pins down the payoff in the following time period which is denoted by $R_{i,t+1} = \alpha^{a_{it}} R^D_{t+1} + (1-\alpha^{a_{it}}) R^I_{t+1}$ where $0 < \alpha^L < \alpha^H < 1$. Because R^I_{t+1} is normalized to 1 for all t, $R_{i,t+1} = 1 + \alpha^{a_{it}}(R^D_{t+1} - 1)$. The total number of IBs whose actions are L is $x_t = \sum_{i=1}^{n} 1_L(a_{it})$ where 1_L is an indicator function whose value is 1 when $a_{it} = L$ and 0 otherwise. R^D_{t+1} is determined by

$$R^D(x_t) = \begin{cases} 1 + \left(\frac{n-x_t}{n}\right) r & \text{for } x_t \in \{0, \ldots, \tau - 1\} \\ 1 + \left(\frac{x_t-n}{n}\right) r & \text{for } x_t \in \{\tau, \ldots, n\} \end{cases} \quad \text{for all } t$$

where r is such that $\max R^D_{t+1} = 1 + r$ with $0 < r < 1$ and $\tau \in \{1, \ldots, n-1\}$ is the threshold at which the coefficient of r changes its sign. This implies that $R_{i,t+1}$ is generated by $R^{a_{it}}(x_t) = 1 + \alpha^{a_{it}}(R^D(x_t) - 1)$.

Coordinated outcome refers to a situation in which x_t is either 0 or n. Here, it is shown that the coordinated outcomes are the only PNEs in IG. This is established by demonstrating that no IB can improve its payoff by changing its action given other IBs' actions only under these outcomes. The impact of an IB's change of action on other IBs' payoffs is also presented for reference.

1. First, consider the case where $x_t = 0$ and $\tau > 1$. When an IB switches its action from H to L while all other IBs' actions are H, the resulting payoff change is

$R^L(1) - R^H(0) = -((\alpha^H - \alpha^L)n + \alpha^L)\frac{r}{n} < 0$, so its payoff worsens. The payoff of other IBs changes by $R^H(1) - R^H(0) = -\alpha^H \frac{r}{n} < 0$, so these IBs' payoffs worsen as well.

2. Next, think about the case where $x_t = 0$ and $\tau = 1$. When an IB switches its action from H to L while all other IBs' actions are H, the resulting payoff change is $R^L(1) - R^H(0) = -(\alpha^L(n-1) + \alpha^H n)\frac{r}{n} < 0$, so its payoff worsens. The payoff of other IBs changes by $R^H(1) - R^H(0) = -\alpha^H(2n-1)\frac{r}{n} < 0$, so these IBs' payoffs also worsen.

3. Next, look into the case where $0 < x_t < \tau$. When an IB switches its action from L to H given other IBs' actions, the resulting payoff change is $R^H(x_t - 1) - R^L(x_t) = ((\alpha^H - \alpha^L)(n - x_t) + \alpha^H)\frac{r}{n} > 0$, so its payoff improves. The payoff of IBs whose current actions are H changes by $R^H(x_t - 1) - R^H(x_t) = \alpha^H \frac{r}{n} > 0$ and the payoff of IBs whose current actions are L changes by $R^L(x_t - 1) - R^L(x_t) = \alpha^L \frac{r}{n} > 0$, so these IBs' payoffs improve too.

4. Next, go over the case where $x_t = \tau$.
   (a) When an IB switches its action from L to H given other IBs' actions, the resulting payoff change is $R^H(\tau - 1) - R^L(\tau) = ((\alpha^H + \alpha^L)(n - \tau) + \alpha^H)\frac{r}{n} > 0$, so its payoff improves. The payoff of IBs whose current actions are H changes by $R^H(\tau - 1) - R^H(\tau) = (2\alpha^H(n - \tau) + \alpha^H)\frac{r}{n} > 0$ and the payoff of IBs whose current actions are L changes by $R^L(\tau - 1) - R^L(\tau) = (2\alpha^L(n - \tau) + \alpha^L)\frac{r}{n} > 0$, so these IBs' payoffs improve as well.
   (b) When an IB switches its action from H to L given other IBs' actions, the resulting payoff change is $R^L(\tau + 1) - R^H(\tau) = ((\alpha^H - \alpha^L)(n - \tau) + \alpha^L)\frac{r}{n} > 0$, so its payoff improves. The payoff of IBs whose current actions are H changes by $R^H(\tau + 1) - R^H(\tau) = \alpha^H \frac{r}{n} > 0$ and the payoff of IBs whose current actions are L changes by $R^L(\tau + 1) - R^L(\tau) = \alpha^L \frac{r}{n} > 0$, so these IBs' payoffs improve too.

5. Next, examine the case where $x_t > \tau$ and $x_t + 1 \leq n$. When an IB switches its action from H to L given other IBs' actions, the resulting payoff change is $R^L(x_t+1) - R^H(x_t) = ((\alpha^H - \alpha^L)(n - x_t) + \alpha^L)\frac{r}{n} > 0$, so its payoff improves. The payoff of IBs whose current actions are H changes by $R^H(x_t + 1) - R^H(x_t) = \alpha^H \frac{r}{n} > 0$ and the payoff of IBs whose current actions are L changes by $R^L(x_t + 1) - R^L(x_t) = \alpha^L \frac{r}{n} > 0$, so these IBs' payoffs also improve.

6. Finally, inspect the case where $x_t = n$. When an IB switches its action from L to H while all other IBs' actions are L, the resulting payoff change is $R^H(n - 1) - R^L(n) = -\alpha^H \frac{r}{n} < 0$, so its payoff worsens. The payoff of other IBs changes by $R^L(n - 1) - R^L(n) = -\alpha^L \frac{r}{n} < 0$, so these IBs' payoffs worsen too.

These 6 cases exhaust all possibilities for x_t. No IB can improve its payoff by changing its action given other IBs' actions only under cases 1 and 2 where $x_t = 0$ and case 6 where $x_t = n$. So, $x_t = 0$ and $x_t = n$ are the only PNEs of IG.

3.6.2 Calibration of the Model

Recall that W_t^a denotes the aggregate wealth in time period t from the actual US real aggregate wealth data (from 1980:Q1 to 2011:Q4) and $W_k^s(\theta)$ the aggregate wealth in time period k generated by the model in this paper based on a vector of parameters $\theta = (\theta_1, \theta_2)$ where $\theta_1 = (n, r, \tau, \varepsilon, \lambda)$ is a sub-vector of parameters which are calibrated to the actual data and $\theta_2 = (\rho, f, \sigma, \alpha^L, \alpha^H, R^I, \beta)$ is a sub-vector of parameters which are fixed at certain values. Because ω in (3.2) is set at $\alpha^H(2n - \tau)\frac{r}{n} + .0001$, it is not included in θ. The calibrated value of θ_1 is

$$\hat{\theta}_1 = \arg\min_{\theta_1} \sum_{l=1}^{6} \left(\frac{m_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a}) - m_l^{BBQ}(\{W_k^s((\theta_1, \theta_2))\}_{k=1}^{T^s})}{s_l^{BBQ}(\{W_t^a\}_{t=1}^{T^a})} \right)^2$$

for a given value of θ_2 where T^a is the length of the actual data and T^s the length of the model-generated data. $T^a = 128$ and $T^s = 500 \times T^a = 64000$.

The numerical minimization is done in two steps: the grid search followed by the amoeba method. The grid search serves two purposes. First, it is used to determine the initial point for the amoeba method. Second, because n and τ are discrete, the grid search is the natural method for pinning down their values. A search over thousands of grid points reveal that $\hat{\theta}_1 = (20, .1, 12, .004, 40)$ is a good place to start the amoeba method which allows a finer search for r, ε, and λ for the given value of n and τ. The amoeba method is chosen because methods that rely on the continuous differentiability of an objective function cannot be used here. It is implemented by FMINSEARCH in Matlab with TolX and TolFun of .01. The final value of $\hat{\theta}_1$ is provided in Table 3.3.

In both the grid search and the amoeba method, first 1000 periods are discarded to reduce the influence of the initial point which is set at $x_t = 0$. This initial point is acceptable because the system can reach the PNEs of IG from an arbitrary initial point under $ITEL$.[44] Numerical experiments show that $\hat{\theta}_1$ is robust to discarding numbers other than first 1000 periods.

---

[44] However, Young (2009) does not discuss how long it takes for this to happen.

3.7 References

1. Acemoglu, D. (2008). "Introduction to modern economic growth," Princeton University Press.

2. Adam, K., Marcet, A., and Nicolini, J. P. (2007). "Stock market volatility and learning," CEPR Discussion Papers 6518, CEPR Discussion Papers.

3. Adda, J. and Cooper, R. W. (2003). "Dynamic economics," MIT Press.

4. Anufriev, M., Assenza, T., Hommes, C., and Massaro, D. (2013). "Interest rate rules and macroeconomic stability under heterogenous expectations," Macroeconomic Dynamics, vol. 17(8), pages 1574-1604, December.

5. Azariadis, C. and Drazen, A. (1990). "Threshold externalities in economic development," The Quarterly Journal of Economics, MIT Press, vol. 105(2), pages 501-26, May.

6. Azariadis, C. and Stachurski, J. (2003). "A forward projection of the cross-country income distribution," KIER Working Papers 570, Kyoto University, Institute of Economic Research.

7. Barro, R. J. and Sala-i-Martin, X. (2003). "Economic growth, 2nd edition," MIT Press.

8. Bergemann, D. and Valimaki, J. (2006). "Bandit problems," Cowles Foundation Discussion Papers 1551, Cowles Foundation for Research in Economics, Yale University.

9. Bertsimas, D. and Tsitsiklis, J. (1993). "Simulated annealing," Statistical Science, vol. 8(1), pages 10-15.

10. Branch, W. A. and Evans, G. W. (2010). "Asset return dynamics and learning," Review of Financial Studies, Oxford University Press for Society for Financial Studies, vol. 23(4), pages 1651-1680, April.

11. _______ (2011). "Learning about risk and return: a simple model of bubbles and crashes," American Economic Journal: Macroeconomics, American Economic Association, vol. 3(3), pages 159-91, July.

12. Brown, J. R., Ivkovic, Z., Smith, P. A., and Weisbenner, S. (2008). "Neighbors matter: causal community effects and stock market participation," Journal of Finance, American Finance Association, vol. 63(3), pages 1509-1531, June.

13. Bry, G. and Boschan, C. (1971). "Programmed selection of cyclical turning points," in Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, pages 7-63, National Bureau of Economic Research, Inc.

14. Camerer, C. and Ho, T. (1999). "Experience-weighted attraction learning in normal form games," Econometrica, Econometric Society, vol. 67(4), pages 827-874, July.

15. Canova, F. (2007). “Methods for applied macroeconomic research,” Princeton University Press.

16. Carceles-Poveda, E. and Giannitsarou, C. (2007). “Adaptive learning in practice,” Journal of Economic Dynamics and Control, Elsevier, vol. 31(8), pages 2659-2697, August.

17. Carvalho, V. M., Martin, A. and Ventura, J. (2012). “Understanding bubbly episodes,” American Economic Review, American Economic Association, vol. 102(3), pages 95-100, May.

18. Cho, I. K. and Sargent, T. J. (2008). “Self-confirming equilibria,” in Durlauf, S. N. and Blume, L. E. (ed.), The New Palgrave Dictionary of Economics. Second Edition. Palgrave Macmillan.

19. Cho, I. K., Williams, N., and Sargent, T. J. (2002). “Escaping Nash inflation,” Review of Economic Studies, Wiley Blackwell, vol. 69(1), pages 1-40, January.

20. Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2009). “Reinforcement learning and savings behavior,” Journal of Finance, American Finance Association, vol. 64(6), pages 2515-2534, December.

21. Cogley, T., Colacito, R., Hansen, L. P., and Sargent, T. J. (2008). “Robustness and U.S. monetary policy experimentation,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 40(8), pages 1599-1623, December.

22. Cogley, T. and Sargent, T. J. (2005). “The conquest of US inflation: learning and robustness to model uncertainty,” Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 8(2), pages 528-563, April.

23. _______ (2008). “Anticipated utility and rational expectations as approximations of Bayesian decision making,” International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 49(1), pages 185-221, February.

24. Cooper, R. W. (1999). “Coordination games,” Cambridge University Press.

25. Dawid, H., Gemkow, S., Harting, P., van der Hoog, S., and Neugart, M. (2012). “Agent-based macroeconomic modeling and policy analysis: the Eurace@Unibi model.”

26. DeJong, D. N. and Dave, C. (2011). “Structural macroeconometrics, 2nd edition,” Princeton University Press.

27. Dixit, A. K. (1990). “Optimization in economic theory,” Oxford University Press.

28. Duffie, D., and Singleton, K. J. (1993). “Simulated moments estimation of Markov models of asset prices,” Econometrica, vol. 61(4), pages 929-952.

29. Dutta, P. (1999). “Strategies and games: theory and practice,” MIT Press.

30. Ellison, M. and Pearlman, J. (2011). “Saddlepath learning,” Journal of Economic Theory, Elsevier, vol. 146(4), pages 1500-1519, July.

31. Ellison, M. and Sargent, T. J. (2012a). “A defense of the Fomc,” International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 53(4), pages 1047-1065, November.

32. _______ (2012b). “Welfare cost of business cycles in economies with individual consumption risk,” Research Discussion Papers 25/2012, Bank of Finland.

33. Ellison, M. and Scott, A. (2013) “Learning and price volatility in duopoly models of resource depletion,” Journal of Monetary Economics, vol. 60(7), October.

34. Ellison, M. and Yates, T. (2007). “Escaping volatile inflation,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 39(4), pages 981-993, June.

35. Erev, I. and Roth, A. E. (1998). “Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria,” American Economic Review, American Economic Association, vol. 88(4), pages 848-81, September.

36. Evans, G. W. and Honkapohja, S. (1999). “Learning dynamics,” in Taylor, J. B. and Woodford, M. (ed.), Handbook of Macroeconomics, Elsevier, vol. 1, chapter 7, pages 449-542.

37. _______ (2001). “Learning and expectations in macroeconomics,” Princeton University Press.

38. _______ (2009). “Learning and macroeconomics,” Annual Review of Economics, Annual Reviews, vol. 1(1), pages 421-451, June.

39. _______ (2011). “Learning as a rational foundation for macroeconomics and finance,” CEPR Discussion Papers 8340, C.E.P.R. Discussion Papers.

40. Fagiolo, G. and Roventini, A. (2012). “Macroeconomic policy in DSGE and agent-based models,” in Gaffard, J. and Napoletano, M. (ed.), Agent-Based Models and Economic Policy, Revue de l’OFCE, Presses de Sciences-Po, pages 67-116.

41. Fernández-Villaverde, J., and Rubio-Ramírez, J. F. (2007). “Estimating macroeconomic models: a likelihood approach,” The Review of Economic Studies, vol. 74(4), pages 1059-1087.

42. Franses, P. H., and Van Dijk, D. (2000). “Non-linear time series models in empirical finance,” Cambridge University Press.

43. Friedman, J. W. (1971). “A non-cooperative equilibrium for supergames,” Review of Economic Studies, Wiley Blackwell, vol. 38(113), pages 1-12, January.

44. Fudenberg, D. and Levine, D. K. (1993). “Self-confirming equilibrium,” Econometrica, Econometric Society, vol. 61(3), pages 523-45, May.

45. _______ (1998). “Learning in games,” European Economic Review, Elsevier, vol. 42(3-5), pages 631-639, May.

46. Gaffard, J. and Napoletano, M. (2012). “Agent-based models and economic policy,” Revue de l’OFCE, Presses de Sciences-Po.

47. Gertler, M. & Kiyotaki, N. (2010). “Financial intermediation and credit policy in business cycle analysis,” in Friedman, B. M. and Woodford, M. (ed.), Handbook of Monetary Economics, Elsevier, vol. 3, chapter 11, pages 547-599.

48. Gilboa, I. and Schmeidler, D. (1989). “Maxmin expected utility with non-unique prior,” Journal of Mathematical Economics, Elsevier, vol. 18(2), pages 141-153, April.

49. Guiso, L., Sapienza, P., and Zingales, L. (2008). “Trusting the stock market,” Journal of Finance, American Finance Association, vol. 63(6), pages 2557-2600, December.

50. Hansen, B. E. (2011). “Threshold autoregression in economics,” Statistics and Its Interface, vol. 4(2), pages 123-128.

51. Hansen, L. P. and Sargent, T. J. (2005). “Robust estimation and control under commitment,” Journal of Economic Theory, Elsevier, vol. 124(2), pages 258-301, October.

52. _______ (2007). “Recursive robust estimation and control without commitment,” Journal of Economic Theory, Elsevier, vol. 136(1), pages 1-27, September.

53. _______ (2008). “Robustness,” Princeton University Press.

54. _______ (2010). “Wanting robustness in macroeconomics,” in Friedman, B. M. and Woodford, M. (ed.), Handbook of Monetary Economics, Elsevier, vol. 3, chapter 20, pages 1097-1157.

55. _______ (2011). “Robustness and ambiguity in continuous time,” Journal of Economic Theory, Elsevier, vol. 146(3), pages 1195-1223, May.

56. Hansen, L. P., Sargent, T. J., Turmuhambetova, G., and Williams, N. (2006). “Robust control and model misspecification,” Journal of Economic Theory, Elsevier, vol. 128(1), pages 45-90, May.

57. Harding, D. and Pagan, A. (2002). “Dissecting the cycle: a methodological investigation,” Journal of Monetary Economics, Elsevier, vol. 49(2), pages 365-381, March.

58. Harley, C. B. (1981). “Learning the evolutionarily stable strategy,” Journal of Theoretical Biology, vol. 89(4), pages 611-633.

59. Hauwe, L. v. d. (2011). “John Maynard Keynes and Ludwig von Mises on probability,” Journal of Libertarian Studies, vol. 22, pages 471-507.

60. Hodrick, R. J. and Prescott, E. C. (1997). “Postwar U.S. business cycles: an empirical investigation,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 29(1), pages 1-16, February.

61. Hong, H., Kubik, J. D., and Stein, J. C. (2004). “Social interaction and stock-market participation,” Journal of Finance, American Finance Association, vol. 59(1), pages 137-163, February.

62. Keynes, J. M. (1921). “Treatise on probability,” MacMillan & Co.

63. _______ (1937). “The general theory of employment,” The Quarterly Journal of Economics, vol. 51(2), pages 209-223.

64. Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). “Optimization by simulated annealing,” Science, vol. 220, no. 4598, pages 671-680, May.

65. Knight, F. H. (1921). “Risk, uncertainty, and profit,” Hart, Schaffner, and Marx.

66. Kreps, D. M. (1998). “Anticipated utility and dynamic choice,” in Frontiers of Research in Economic Theory: the Nancy L. Schwartz Memorial Lectures, 1983-1997, Cambridge University Press.

67. Lee, B. S. and Ingram, B. F. (1991). “Simulation estimation of time-series models,” Journal of Econometrics, Elsevier, vol. 47(2-3), pages 197-205, February.

68. Lengnick, M. (2011). “Agent-based macroeconomics - a baseline model,” Economics Working Paper 2011, Christian-Albrechts-Universität Kiel, Department of Economics.

69. McLaren, J. (2003). “Trade and market thickness: effects on organizations,” Journal of the European Economic Association, MIT Press, vol. 1(2-3), pages 328-336.

70. Miao, J. (2004). “A note on consumption and savings under Knightian uncertainty,” Annals of Economics and Finance, Society for AEF, vol. 5(2), pages 299-311, November.

71. Mises, L. v. (1940). “Nationalökonomie: theorie des handelns und wirtschaftens,” Wirtschaft und Finanzen.

72. _______ (1949). “Human action: a treatise on economics,” Yale University Press.

73. Nax, H. H., Burton-Chellew, M. N., West, S. A., and Young, H. P. (2013). “Learning in a black box,” Department of Economics Discussion Paper Series 653, University of Oxford, April.

74. Pradelski, B. S. and Young, H. P. (2012). “Learning efficient Nash equilibria in distributed systems,” Games and Economic Behavior, Elsevier, vol. 75(2), pages 882-897, July.

75. Roth, A. E. and Erev, I. (1995). “Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term,” Games and Economic Behavior, Elsevier, vol. 8(1), pages 164-212.

76. Roth, A. E. (2009). “What have we learned from market design?,” in Innovation Policy and the Economy, National Bureau of Economic Research, Inc., vol. 9, pages 79-112.

77. Rutenbar, R. A. (1989). “Simulated annealing algorithms: an overview,” Circuits and Devices Magazine, IEEE, vol. 5(1), pages 19-26.

78. Sargent, T. J. (1999). “The conquest of American inflation,” Princeton University Press.

79. _______ (2008). “Evolution and intelligent design,” American Economic Review, American Economic Association, vol. 98(1), pages 5-37, March.

80. Sargent, T. J., Williams, N., and Zha, T. (2006). “Shocks and government beliefs: the rise and fall of American inflation,” American Economic Review, American Economic Association, vol. 96(4), pages 1193-1224, September.

81. _______ (2009). “The conquest of South American inflation,” Journal of Political Economy, University of Chicago Press, vol. 117(2), pages 211-256, April.

82. Song, R., Jang, S., Hanssens, D., and Suh, J. (2013). “When overconfidence meets reinforcement learning,” Working Paper.

83. Stigler, S. M. (2010). “The changing history of robustness,” The American Statistician, American Statistical Society, vol. 64(4), pages 277-281.

84. Stiglitz, J. E. (1994). “The role of the state in financial markets,” in Proceedings of the World Bank Annual Conference on Development Economics 1993, The World Bank.

85. Taylor, J. B. (1993). “Discretion versus policy rules in practice,” Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 39(1), pages 195-214, December.

86. Williams, N. (2002). “Escape dynamics in learning models,” Working Paper.

87. _______ (2008). “Robust control,” in Durlauf, S. N. and Blume, L. E. (ed.), The New Palgrave Dictionary of Economics. Second Edition. Palgrave Macmillan.

88. _______ (2009). “Escape dynamics in learning models,” Working Paper.

89. Young, H. P. (2009). “Learning by trial and error,” Games and Economic Behavior, Elsevier, vol. 65(2), pages 626-643, March.

Chapter 4

Portfolio Choice under Knightian Uncertainty: Application of Interactive Trial and Error Learning

Abstract

Portfolio choice under Knightian Uncertainty is modeled using Interactive Trial and Error Learning (Pradelski and Young, 2012). The portfolio choice model can match the first two moments of the US real excess return of equity over bonds almost perfectly. It can also reproduce non-linearity in the data at a non-negligible frequency.

4.1 Introduction

This paper presents another application of Interactive Trial and Error Learning ($ITEL$) (Young, 2009; Pradelski and Young, 2012) in the context of macroeconomics. Again, the model agents face Knightian Uncertainty (Knight, 1921) because the structure of the economic system is unknown to them. Whereas the previous chapter was concerned with how Knightian Uncertainty about the return on investment affects other aggregate variables, this paper deals with Knightian Uncertainty about the return on investment directly. Here, the center of attention is portfolio choice by $ITEL$. The portfolio choice model is calibrated against the real excess return of equity over bonds in the US financial market. It will be shown that the model can match several features of the actual data reasonably well. These include the mean (i.e., equity premium) and the variance as well as non-linearity in the form of threshold effect. The threshold effect here means that the relationship between the current excess return and the past excess return changes depending on the value of the past excess return. Because the previous chapter provides the relevant literature review about model uncertainty in macroeconomics, it is omitted here. The modeling strategy in this paper contributes to a large literature on equity premium puzzles. Mehra and Prescott (2003) provides a list of approaches that have been considered in the literature.

The paper is structured as follows: section 4.2 analyzes the US real excess return data; section 4.3 presents the model; section 4.4 discusses calibration and simulation; section 4.5 examines whether the agents have incentives to adopt $ITEL$; section 4.6 concludes.

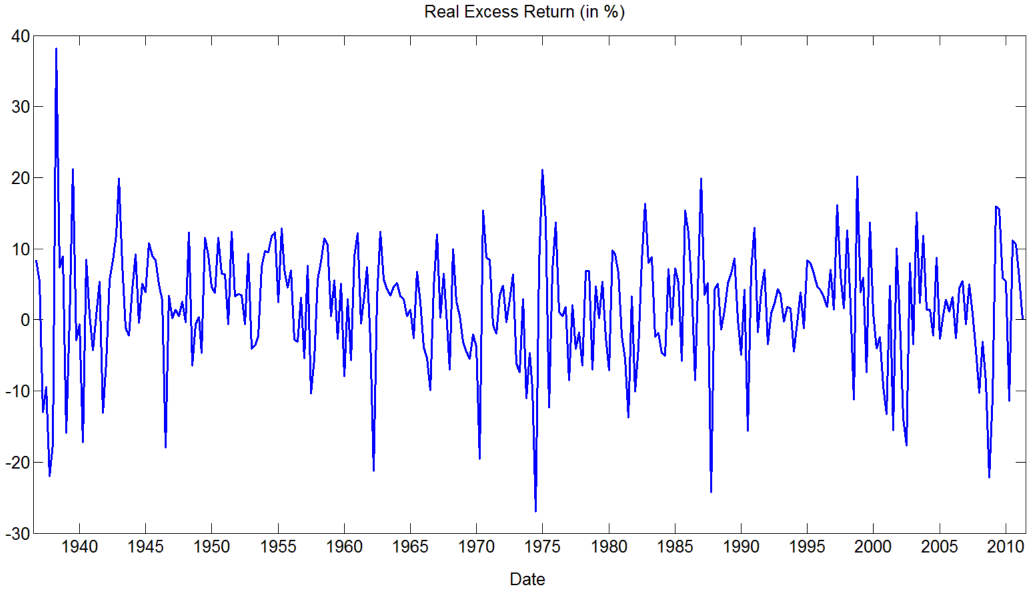
4.2 US Real Excess Return Data

Let us start this section by taking a look at the real excess return of equity over bond (RER henceforth) in the US financial market. Figure 4.1 plots RER from 1936:Q4 to 2011:Q2[1]. One can see that it is characterized by large fluctuations (the range of

---

[1] It is computed using S&P 500 Total Return Index, USA 90 Day T-bill Secondary Market Rate, and USA BLS Consumer Price Index which are retrieved from "The Global Financial Database

Figure 4.1: Real Excess Return (1936:Q4 to 2011:Q2)



The number along the horizontal axis corresponds to the first quarter of the year that the number refers to.

approximately 65%) and these fluctuations recur over time with no sign of decline. In what follows, it will be subject to more rigorous statistical analysis.

4.2.1 Summary Statistics

Table 4.1 provides summary statistics for RER. The first four rows contain the first four moments of RER based on the entire sample. The mean of 1.8125% indicates the equity premium while the variance of 70.3893%^{2} reveals a very volatile nature of RER. The skewness of -.3171% suggests that a larger mass of the data points are located on the right side of the mean while the kurtosis of 4.5990% points out that the distribution of RER has fatter tails than a normal distribution. Figure 4.2 fits the distribution of RER against a normal distribution whose mean and variance are equal to the values above. It shows that the fit is poor around the mean and near the tails of the distribution. Jarque-Bera test (see Jarque and Bera, 1987) also strongly

---

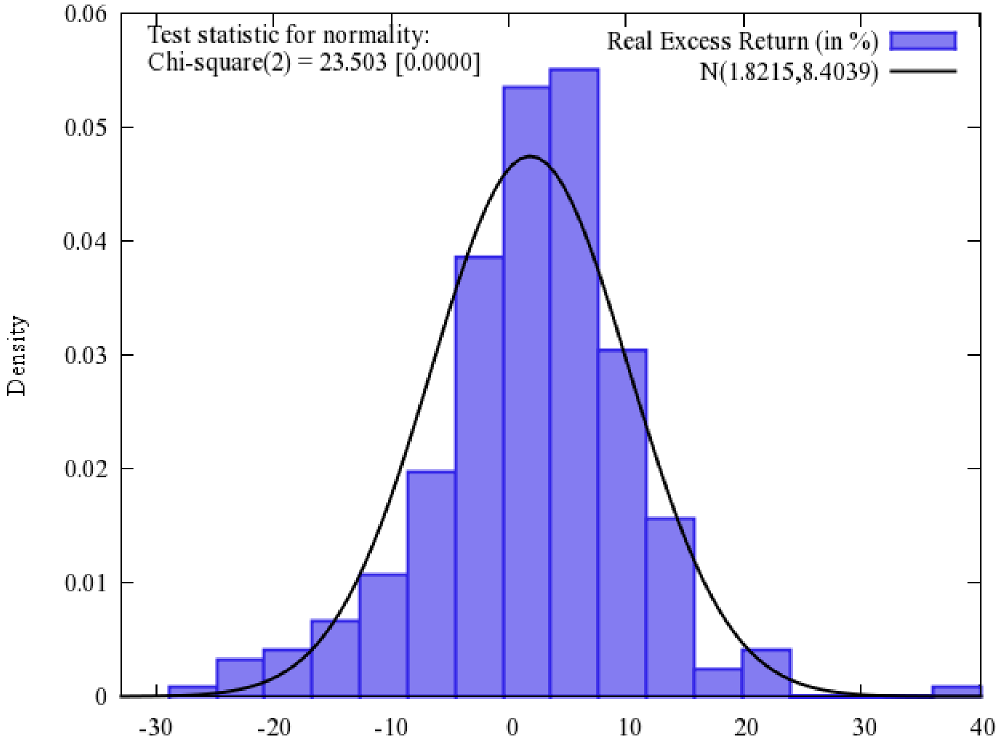
(http://www.globalfinancialdata.com).” The computation is based on Adam, Marcet, and Nicolini (2011) but the resulting series is truncated below due to (a) the availability of dividend data set (available from 1936:Q1 onward and contained in S&P 500 Total Return Index) and (b) the seasonal adjustment of dividends by taking the unweighted average of the current dividend and dividends from the past three quarters (see Campbell, 2003).

Table 4.1: Summary Statistics for Real Excess Return

Statistic	Value
Mean	1.8215%
Variance	70.3893%^{2}
Skewness	-0.3171%
Kurtosis	4.5990%
Mean$^+$	2.5040%
Variance$^+$	54.1413%^{2}
Skewness$^+$	-0.3877%
Kurtosis$^+$	3.9055%
Mean$^-$	0.5598%
Variance$^-$	96.8132%^{2}
Skewness$^-$	-0.0844%
Kurtosis$^-$	4.4906%

The table provides summary statistics for RER. The entries in the first four rows are based on the entire sample. The entries in the next four rows are based on a sub-sample whose elements are data points for which the value of RER in the previous quarter is greater than zero. The entries in the last four rows are based on a sub-sample whose elements are data points for which the value of RER in the previous quarter is less than or equal to zero.

Figure 4.2: Fit of Real Excess Return to a Normal Distribution



rejects the null hypothesis of normality as indicated by the $p - value$ on the figure. This non-normality of RER is conventional wisdom (see Campbell, Lo, and Mackinlay, 1997).

To take a closer look at the asymmetric nature of the RER distribution (as evidenced by its skewness), the data are split into two sub-samples depending on whether the value of RER in the previous quarter is greater than zero or not. The entries in the fifth to the eighth row of Table 4.1 correspond to the first four moments conditional on the value of RER in the previous quarter being positive (denoted by superscript $+$) and the entries in the remaining rows to the first four moments conditional on the value of RER in the previous quarter being less than or equal to zero (denoted by superscript $-$). They suggest that the behavior of RER depends on which side of zero it belongs to in the previous quarter. For instance, while there seems to be a positive relationship between the current RER and the previous quarter's RER when the latter is positive ($\text{Mean}^{+} = 2.5040\%$), the opposite seems to be the case when it is non-positive ($\text{Mean}^{-} = .5598\%$). Also, RER is more volatile and outlier-prone conditional on the previous quarter's RER being non-positive than it being positive as indicated by the variances and the kurtoses. The next subsection analyzes these observations more systematically by taking the data to a threshold model.

4.2.2 Threshold Effect

The observations in the previous subsection can be interpreted as an outcome of threshold effect with zero being the threshold at which the behavior of RER changes. In order to ascertain the existence of the threshold effect and pin down the location of the threshold more rigorously, Self-Exciting Threshold Autoregressive ($SETAR$) models are fitted to the data. $SETAR$ models are one of the standard methods for studying threshold effects (see Hansen, 1997; 1999) along with other methods such as Smooth Transition Autoregressive ($STAR$) models and Markov-Switching models (see Franses and van Dijk, 2000). They are chosen over other methods because they bring out a relationship between the current and the past RERs more explicitly. Let us start by taking a $SETAR$ model with 2 regimes and the lag order of 12 (in both regimes) or $SETAR(2, 12)$ to the data. Formally, for $t \geq 1$,

$$RER_t = (\phi_0^1 + \sum_{j=1}^{12} \phi_j^1 RER_{t-j}) \times 1(\underbrace{RER_{t-1} \leq \gamma}_{regime\,1}) + (\phi_0^2 + \sum_{j=1}^{12} \phi_j^2 RER_{t-j}) \times 1(\underbrace{RER_{t-1} > \gamma}_{regime\,2}) + e_t, \tag{4.1}$$

where t is time subscript, ϕ's are the autoregressive parameters, superscripts on ϕ's indicate the regime, subscripts on ϕ's indicate the lag order, $1(.)$ is an indicator function whose value is 1 if the statement inside the parentheses is true and 0 otherwise, γ is the threshold parameter, and e_t is an *iid* $(0,\sigma^2)$ residual term. This representation makes clear that RER_t's behavior changes depending on the value of RER_{t-1}.

Table 4.2 presents the estimates $\hat{\gamma}$ and $\hat{\phi}$'s with the latter's levels of statistical significance indicated by superscripts * (10%), ** (5%), and *** (1%).[2] Newey-West standard errors are used for computing the corresponding $p-values$ in order to make adjustments for possible heteroscedasticity and autocorrelation (see Newey and West, 1987). Each regime is constrained to contain at least 10% of the observations following Hansen (1999). The result indicates a negative relationship between RER_t and RER_{t-1} in regime 1 but a positive relationship between the two in regime 2 as conjectured in subsection 4.2.1. However, the level of statistical significance in regime 1 is only at 10% level. In contrast, the level of statistical significance in regime 2 is at 1% level. Appendix 4.7 discusses how $\hat{\gamma}$ and $\hat{\phi}$'s are obtained in detail. Figure 4.3 shows the fit of $SETAR(2, 12)$ to the actual RER. While the model tracks the direction of the movements in RER reasonably well, it does not get the magnitude right.

To see whether inserting an extra regime can bring about a negative relationship between RER_t and RER_{t-1} at a higher level of statistical significance, a $SETAR$ model with 3 regimes and the lag order of 12 (in all regimes) or $SETAR(3, 12)$ is taken to the data this time. The only difference from $SETAR(2, 12)$ above is that more parameters need to be estimated now (see Hansen, 1999). In particular, two

---

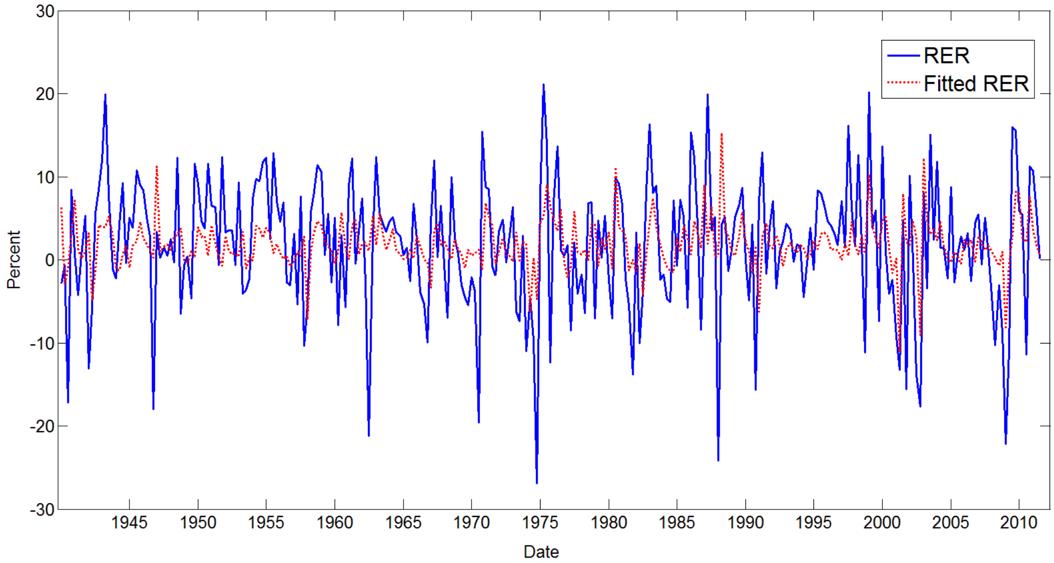
[2] Even though $\hat{\gamma}$ has non-standard asymptotic properties, it is possible to construct a confidence interval for $\hat{\gamma}$ based on bootstrapping (see Hansen, 1997). But this is beyond the scope of the analysis here.

Table 4.2: $SETAR(2,12)$

	$RER_{t-1} \leq -7.0863$ Estimate	$RER_{t-1} > -7.0863$ Estimate
$constant$	-7.0017	1.3556^{*}
RER_{t-1}	-0.5748^{*}	0.2482^{***}
RER_{t-2}	-0.4152^{*}	-0.0248
RER_{t-3}	0.5605^{**}	-0.0609
RER_{t-4}	0.5535^{**}	-0.0631
RER_{t-5}	-0.0091	0.0085
RER_{t-6}	0.0828	-0.0411
RER_{t-7}	-0.0395	-0.0892^{*}
RER_{t-8}	0.6941	0.0231
RER_{t-9}	-0.5999^{***}	-0.0222
RER_{t-10}	-0.0660	0.0447
RER_{t-11}	-0.4372^{*}	0.0047
RER_{t-12}	0.1451	0.0084

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Figure 4.3: Fit of $SETAR(2,12)$



The number along the horizontal axis corresponds to the first quarter of the year that the number refers to. The straight line plots the actual RER and the dotted line the fitted RER.

threshold parameters, γ_1 and γ_2, need to be estimated for three regimes $RER_{t-1} \leq \gamma_1$, $\gamma_1 < RER_{t-1} \leq \gamma_2$, and $RER_{t-1} > \gamma_2$ (regimes 1, 2, and 3 respectively). Again, Newey-West standard errors are used to compute the $p-values$ for the estimates of ϕ's (which are the autoregressive parameters). The restriction that each regime contains at least 10% of the observations is imposed here as well.

Table 4.3 gives the estimates $\hat{\gamma}_1$, $\hat{\gamma}_2$, and $\hat{\phi}$'s. Note that while regime 1 of $SETAR(3,12)$ is identical to regime 1 of $SETAR(2,12)$, regime 2 of $SETAR(2,12)$ is split into two to form regimes 2 and 3 of $SETAR(3,12)$. The table answers the question above by reporting that the negative relationship between RER_t and RER_{t-1} in regime 2 is statistically significant at 1% level. The positive relationship between RER_t and RER_{t-1} in regime 3 is also statistically significant at 1% level. Figure 4.4 shows that the fit of $SETAR(3,12)$ to the actual RER is also better than the fit of $SETAR(2,12)$.[3] The result from $SETAR(3,12)$ supports the conjecture that there is a reversal from the positive relationship between RER_t and RER_{t-1} to the negative relationship across the different regimes.

4.2.3 Discussion

Given that the standard macroeconomic models have hard time rationalizing the equity premium (see Mehra and Prescott, 1985; 2003), which is about the first moment of RER, it is highly unlikely that they can account for the higher moments of RER. Moreover, it is highly unlikely that the standard macroeconomic models can produce the threshold effect documented in subsection 4.2.2 which should be responsible for the behavior of RER (including all the moments). The aim of the following subsections is to construct a model that comes closer to matching the empirical findings of this section.

---

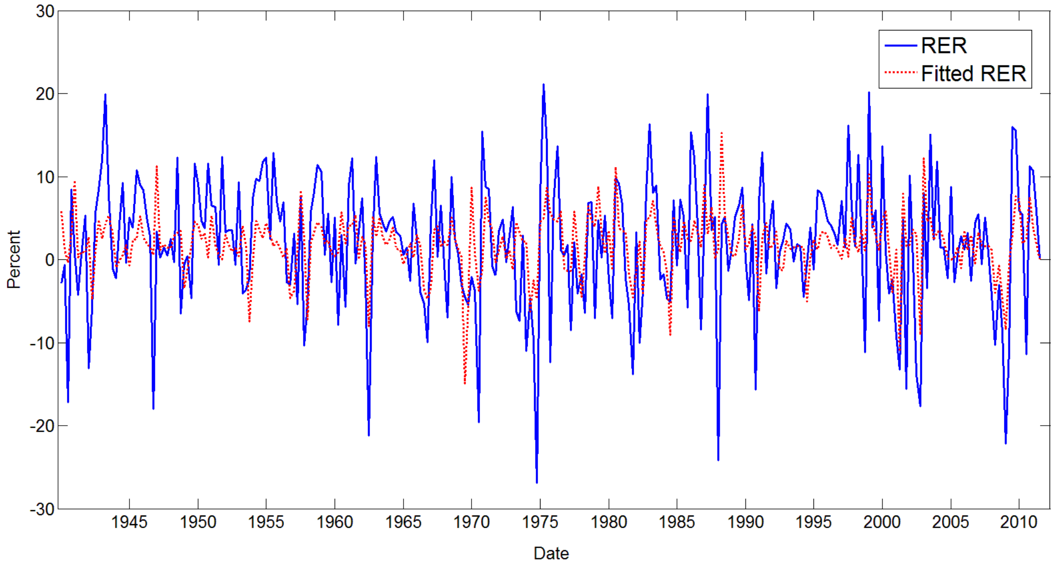
[3] It is possible to do a specification test to choose between $SETAR(2,12)$ and $SETAR(3,12)$ (see Hansen, 1999). However, this is beyond the scope of the analysis here.

Table 4.3: $SETAR(3,12)$

	$RER_{t-1} \leq -7.0863$ Estimate	$-7.0863 < RER_{t-1} \leq -2.7500$ Estimate	$RER_{t-1} > -2.7500$ Estimate
$constant$	-7.0017	-17.2346^{***}	1.3140
RER_{t-1}	-0.5748^{*}	-3.2287^{***}	0.2546^{***}
RER_{t-2}	-0.4152^{*}	0.0581	-0.0342
RER_{t-3}	0.5605^{**}	-0.5490^{**}	-0.0508
RER_{t-4}	0.5535^{**}	-0.4985^{***}	-0.0358
RER_{t-5}	-0.0091	0.5003^{*}	0.0126
RER_{t-6}	0.0828	0.0488	-0.0347
RER_{t-7}	-0.0395	0.0767	-0.0909^{*}
RER_{t-8}	0.6941	-0.0401	0.0460
RER_{t-9}	-0.5999^{***}	0.4850^{***}	-0.0480
RER_{t-10}	-0.0660	-0.3093^{***}	0.0927^{*}
RER_{t-11}	-0.4372^{*}	0.1486	-0.0120
RER_{t-12}	0.1451	0.2307	0.0066

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Figure 4.4: Fit of $SETAR(3,12)$



The labeling conventions are identical to those in Figure 4.3.

4.3 Model

4.3.1 Setup

Unlike the model in the previous chapter where households were divided into bankers and depositors, the model in this paper features bankers only as it is not concerned with other aggregate variables such as the aggregate wealth. Again, the economy is divided into n islands each of which is occupied by a continuum of infinitely-living households of unit measure. The households on each island invest their endowments in every time period and consume the fruits of their investments in the following time period. They do not have access to a reliable storage technology so they have to consume the entire fruits. For simplicity, the size of the endowment is normalized to 1 and fixed over time for each household. As before, the households on each island face Knightian Uncertainty (Knight, 1921) which is due to their lack of knowledge about the structure of the economic system: they know nothing about other islands and do not know how the return on investment is generated. However, there is perfect information within the same island.

Because this paper is not concerned with the risk-attitude of the households, it is assumed that the households' preference is linear in consumption: $u(c) = c$ where $c \geq 0$ is a consumption. Let y denote the gross return on investment. Given the strictly monotone preference for consumption and the absence of a reliable storage technology, it follows that $c = y > 0$ whenever $y > 0$.

The households on each island, which are banker-consumers, are modeled with the representative household and is referred to as the Island Banker (IB) so that the terminology in this paper is consistent with the one in the previous chapter.

4.3.2 Investment

Again, two types of assets are available for investment: the independent (I) and the dependent (D). IB on island i invests its endowment in each time period by choosing a portfolio from $A_i = \{.1, .2, \ldots, .9, 1\}$ (a set whose element increases by .1 from .1

until 1) where A_i is a set of portfolios whose elements correspond to proportions of the endowment invested in D.[4] A_i is invariant across i. As before, the gross rate of return on I in time period $t+1$, which is denoted by R^I_{t+1}, is fixed over time and normalized to 1: $R^I_{t+1} = R^I = 1$ for all t. The gross rate of return on D in time period $t+1$, which is denoted by R^D_{t+1}, comes with the lower bound of $\underline{R}^D$ and the upper bound of $\overline{R}^D$. $0 < \underline{R}^D < R^I < \overline{R}^D$ so that neither D nor I is superior at all times. Let $a_{it} \in A_i$ be a portfolio chosen by IB on island i in time period t and a_t a n-by-1 vector whose i_{th} element is a_{it}. R^D_{t+1} is generated by a mechanism

$$R^D(a_t) = \underline{R}^D + (\overline{R}^D - \underline{R}^D) \times \kappa(a_t) \tag{4.2}$$

where $0 \leq \kappa(.) \leq 1$ and has the following properties:

1. $\forall i$, $a_{it} \geq a^{'}_{it}$ and $a_{jt} \geq a^{'}_{jt}$ $\forall j \neq i \Rightarrow \kappa(a_t) \leq \kappa(a^{'}_t)$;
2. $\forall i$, $a_{it} > a^{'}_{it}$ and $a_{jt} \geq a^{'}_{jt}$ $\forall j \neq i \Rightarrow \kappa(a_t) < \kappa(a^{'}_t)$.

They imply that i's choice has negative externalities on others. Taking I as government debts, D as equities, and a_t as a measure of market thickness for D as in the previous chapter, this way of modeling R^D can be rationalized on the grounds that (a) there is a positive relationship between market thickness and market liquidity[5] and (b) there is a negative relationship between market liquidity and excess return of equity over bond[6]. Here, $\kappa(.)$ takes the functional form of

$$\kappa(a_t) = \left(1 - \frac{\sum_{i=1}^{n} a_{it}}{n}\right)^{\eta} \tag{4.3}$$

with $\eta \geq 0$. The gross return on investment for IB on island i in time period $t+1$ is

$$y_{i,t+1} = R^I(1 - a_{it}) + R^D_{t+1} a_{it}. \tag{4.4}$$

---

[4] The non-inclusion of 0 in A_i implies that IBs cannot divest from D fully.

[5] See Lippman and McCall(1986).

[6] For instance, see Amihud (2002).

$y_{i,t+1} > 0$ because $0 < \underline{R}^D < R^I < \overline{R}^D$. Henceforth, a_{it} is referred to as action, a_t as action profile, and $y_{i,t+1}$ as payoff.

4.3.3 Interactive Trial and Error Learning

As in the previous chapter, the Island Bankers (IBs) choose their investment portfolios by following Interactive Trial and Error Learning ($ITEL$) in order to cope with Knightian Uncertainty. The version of $ITEL$ being used here is Pradelski and Young (2012) because it fits the data better.[7] The difference between Young (2009) (featured in subsubsection 3.3.2.2) and Pradelski and Young (2012) in terms of learning rules is that whereas a new action that yields a higher payoff than the current benchmark payoff is always adopted under the active search in the former, it is adopted with a probability that is increasing in the payoff gain in the latter. The difference between Young (2009) and Pradelski and Young (2012) in terms of properties is that whereas the former implements a Pure Nash Equilibrium of a game a very high proportion of the time as the probability of the active search becomes smaller, the latter implements an "efficient" or Pareto-dominant Pure Nash Equilibrium. For the ease of reference, $ITEL$ of Pradelski and Young (2012) is presented below.

Recall that the state vector of IB on island i in time period $t-1$ is $z_{i,t-1} = (m_{i,t-1}, \overline{a}_{i,t-1}, \overline{y}_{i,t-1})$ where m is the mood (content (c), discontent (d), hopeful (h), and watchful (w)), $\overline{a}$ is the benchmark action, and $\overline{y}$ is the benchmark payoff. The learning rules are as follows:

- $m_{i,t-1} = c$: In $t-1$, IB on island i experiments with an action that is different from the benchmark action with a probability of ε (active search). If the experiment pushes through, a new action is drawn from $A_i - \{\bar{a}_{i,t-1}\}$ uniformly at random.

    - $a_{i,t-1} \neq \bar{a}_{i,t-1}$ and $y_{it} \leq \bar{y}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $a_{i,t-1} \neq \bar{a}_{i,t-1}$ and $y_{it} > \bar{y}_{i,t-1} \Rightarrow$

---

[7] This is because Pradelski and Young (2012) has decision rules with more random components which make them more compatible with lower frequency data like RERs.

        - $m_{it} = c$, $\overline{a}_{it} = a_{i,t-1}$, and $\overline{y}_{it} = y_{it}$ with a probability $\varepsilon^{G(y_{it}-\overline{y}_{i,t-1})}$ (G is defined below);
        - $m_{it} = c$, $\overline{a}_{it} = \overline{a}_{i,t-1}$, and $\overline{y}_{it} = \overline{y}_{i,t-1}$ with a probability $1 - \varepsilon^{G(y_{it}-\overline{y}_{i,t-1})}$.
    - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $y_{it} = \bar{y}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $y_{it} > \bar{y}_{i,t-1} \Rightarrow m_{it} = h$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $a_{i,t-1} = \bar{a}_{i,t-1}$ and $y_{it} < \bar{y}_{i,t-1} \Rightarrow m_{it} = w$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.

- $m_{i,t-1} = h$: In $t-1$, IB on island i maintains its benchmark action.
    - $y_{it} = \bar{y}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $y_{it} > \bar{y}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = y_{it}$.
    - $y_{it} < \bar{y}_{i,t-1} \Rightarrow m_{it} = w$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.

- $m_{i,t-1} = w$: In $t-1$, IB on island i maintains its benchmark action.
    - $y_{it} = \bar{y}_{i,t-1} \Rightarrow m_{it} = c$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $y_{it} > \bar{y}_{i,t-1} \Rightarrow m_{it} = h$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.
    - $y_{it} < \bar{y}_{i,t-1} \Rightarrow m_{it} = d$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$.

- $m_{i,t-1} = d$: In $t-1$, IB on island i draws $a_{i,t-1}$ from A_i uniformly at random (passive search).
    - $m_{it} = c$, $\bar{a}_{it} = a_{i,t-1}$, and $\bar{y}_{it} = y_{it}$ with a probability of $\varepsilon^{F(y_{it})}$ (F is defined below).
    - $m_{it} = d$, $\bar{a}_{it} = \bar{a}_{i,t-1}$, and $\bar{y}_{it} = \bar{y}_{i,t-1}$ with a probability of $1 - \varepsilon^{F(y_{it})}$.

There are some restrictions on F and G. $0 < F < \frac{1}{2n}$ and $0 < G < \frac{1}{2}$ ensure that ε^F and ε^G are not small. The functional form restrictions of

$$F = -\varphi_1 y_{it} + \varphi_2 \tag{4.5}$$

and

$$G = -\delta_1(y_{it} - \overline{y}_{i,t-1}) + \delta_2, \tag{4.6}$$

where $\varphi_1, \delta_1 > 0$ make ε^F and ε^G strictly increasing in y_{it} and $y_{it} - \overline{y}_{i,t-1}$ respectively. To make sure that these hold, $\varphi_1 = \frac{1}{2n(\overline{R}^D - \underline{R}^D)}$, $\varphi_2 = \frac{\overline{R}^D}{2n(\overline{R}^D - \underline{R}^D)}$, $\delta_1 = \frac{1}{2(\overline{R}^D - \underline{R}^D)}$, and $\delta_2 = \frac{1}{2}$.

4.4 Calibration and Simulation

4.4.1 Calibration

Let RER^a_t denote the real excess return in time period t from the actual data (from 1936:Q4 to 2011:Q2) and $RER^s_k(\theta)$ the real excess return in time period k generated by the model in this paper based on a vector of parameters θ. Here, $RER^s_k = (R^D_k - R^I)/R^I \times 100\%$. Recall that the model in this paper contains parameters n, R^I, $\underline{R}^D$, $\overline{R}^D$, η, and ε. The restrictions that $\underline{R}^D = R^I - \rho$ and $\overline{R}^D = R^I + \rho$ are imposed in order to reduce the number of parameters. As mentioned in subsection 4.3.2, $R^I = 1$. So, $\theta = (\theta_1, \theta_2)$ where $\theta_1 = (n, \varepsilon, \rho, \eta)$ is calibrated to the actual data and $\theta_2 = R^I$ is fixed at 1. Because calibrating ρ, η, and ε jointly leads to a convergence problem (as they perform similar roles in the model to a certain extent), the calibration proceeds by optimizing ρ and η for given values of n and ε and selecting the best pair of n and ε in the end.

As in the previous chapter (refer to subsection 3.4.1), the model in this paper is calibrated by the Simulated Method of Moments. The moments to be matched are the mean and the variance of RER from the actual data as the vast majority of empirical researchers focus on these. Let h be a function that maps the data (actual and model-generated) into the components of the moments. For X_t,

$$h(X_t) = \begin{pmatrix} X_t \\ (X_t - \bar{X})^2 \end{pmatrix}$$

Table 4.4: Parameter Values

Parameter	Value
n	10
ε	.4
R^I	1
ρ	.5346
η	.8444

where $\bar{X}$ is the sample mean. For given values of n and ε, ρ and η are chosen by

$$(\tilde{\rho}, \tilde{\eta}) = \arg\min_{\rho,\eta} \sum_{l=1}^{2} \left(\frac{1}{T^a} \sum_{t=1}^{T^a} h_l(RER_t^a) - \frac{1}{T^s} \sum_{k=1001}^{T^s+1000} h_l(RER_k^s(\theta_1, \theta_2)) \right)^2 \qquad (4.7)$$

where T^a is the length of the actual data and T^s is the length of the model-generated data. Here, $T^a = 299$ and $T^s = 100 \times T = 29900$. k starts at 1001 because the initial 1000 observations are deleted in order to reduce their influence on the calibration.[8] $(\hat{n}, \hat{\varepsilon})$ is the pair that minimizes the quadratic objective function in (4.7) and $(\hat{\rho}, \hat{\eta}) = (\tilde{\rho}(\hat{n}, \hat{\varepsilon}), \tilde{\eta}(\hat{n}, \hat{\varepsilon}))$ is the pair that corresponds to $(\hat{n}, \hat{\varepsilon})$. As discussed in subsection 3.4.1, this is not estimation as the model in this section is not differentiable with respect to its parameters.[9] Table 4.4 provides the parameter values. For $\hat{\theta} = (\hat{\theta}_1, \theta_2)$, the quadratic objective function in (4.7) converges to $.00006905\%^2$ which is practically zero. This implies that the model in this paper can account for the mean and the variance of RER from the actual data almost perfectly. $\varepsilon = .4$ in the table indicates that Interactive Trial and Error Learning ($ITEL$) can be empirically relevant even when the probability of the active search is large (even though the optimality property discussed in subsection 4.3.3 no longer applies).

4.4.2 Simulation

Table 4.5 gives some summary statistics for the model-generated RER data (under "Model"). The reported numbers are averages of 500 simulation rounds. The length

---

[8] The calibration is not sensitive to how the initial conditions are set. Here, they are (a) $\forall i$, set $m_{i2} = c$, (b) $\forall i$, draw $\overline{a}_{i1}$ from A_i uniformly at random, (c) $\forall i$, set $a_{i1} = \overline{a}_{i1}$ and $\overline{a}_{i2} = \overline{a}_{i1}$, and (d) $\forall i$, compute the corresponding y_{i2} and set $\overline{y}_{i2} = y_{i2}$ and $\overline{y}_{i1} = \overline{y}_{i2}$.

[9] Thus, it is not concerned with issues that are more pertinent to estimation such as the identification problem.

Table 4.5: Some Summary Statistics

Statistic	Model	Actual
Mean (%)	1.8425	1.8215
Variance ($\%^2$)	69.8336	70.3893
Skewness (%)	-0.2026	-0.3171
Kurtosis (%)	3.0770	4.5990

The entries under Model are based on the model-generated RER data (averages of 500 simulation rounds) and the entries under Actual based on the actual RER data.

of each round is set to the length of the actual data (299 periods) after deleting first 1000 periods in order to reduce the influence of the initial conditions. The model is parameterized at the values in Table 4.4. As expected from how the calibration was conducted in the previous subsubsection, Table 4.5 shows that the mean and the variance from the model-generated data are very close to the ones from the actual data. The skewness from the model-generated data is fairly close to the one from the actual data as well even though it was not involved in the calibration. The kurtosis from the model-generated data is too small compared to the one from the actual data. However, the kurtosis in excess of 4 does happen and the kurtosis in excess of 3.5 happens rather frequently (11.4% of the time).

How frequently does the model produce the threshold effect documented in subsection 4.2.2? In order to answer this question, let us focus on the statistically significant negative relationship between RER_t and RER_{t-1} in regime 2 and the statistically significant positive relationship between RER_t and RER_{t-1} in regime 3 of $SETAR(3, 12)$. In the model-generated data (which were used to produce Table 4.5), these occur (jointly) 4.4% of the time at 1% level of statistical significance and 6.6% of the time at 5% level of statistical significance. Thus, the model can generate the threshold effect at non-negligible frequency even though it was not involved in the calibration. The frequency of the threshold effect can increase to 6.2% at 1% level of statistical significance and 9.8% at 5% level of statistical significance when ρ and η are slightly changed to .5338 and .8471 respectively. The mean and the variance are 1.8986% and $70.4890\%^2$ respectively in this case (averages of 500 simulation rounds) which are also close to the ones based on the actual data.

Overall, the model is reasonably consistent with the stylized facts of the actual RER data.

4.5 Experiment

Suppose one of the Island Bankers (IBs), say the one on island g, operates under Rational Expectations. By how much does this alter the behavior of RER^s (the model-generated RER)? Now, IB on island g solves

$$\max_{a_{g,t-1}\in A_g} y_{gt} = R^I(1-a_{g,t-1})+R_t^D a_{g,t-1}\; s.t.\; R_t^D = \underline{R}^D+(\overline{R}^D-\underline{R}^D)\times\left(1-\frac{a_{g,t-1}+\sum_{j\neq g} a_{j,t-1}}{n}\right)^{\eta}$$

in time period $t-1$ $\forall t$. To answer the question above, this modified model is simulated 500 times under the same conditions as subsection 4.4.2 except for the introduction of the Rational Expectations IB. Table 4.6 gives some summary statistics which are averages of 500 simulation rounds. The introduction of the Rational Expectations IB increases the mean of RER^s only modestly by .26% but decreases the variance dramatically by 40.86%[2]. It also makes RER^s more negatively skewed and more outlier-prone.

One may think of IB on island g as a sophisticated investor which has better information about the state of the financial market than other IBs. Given the observed improvements in the mean and the variance of RER^s, one should ask whether the introduction of this investor leads to significant improvements in the payoffs. The simulation results (based on the same model-generated data as above) indicate that while the Rational Expectations IB's per period utility increases by 1.68% relative to when it follows $ITEL$, other IBs' per period utility increases by .39% only. Therefore, the introduction of the Rational Expectations IB does not lead to significant payoff improvements. The higher the cost of maintaining Rational Expectations (e.g., collection of the relevant data), the weaker the incentive of IB on island g to give up $ITEL$.

Table 4.6: Some Summary Statistics

Statistic	Value
Mean	2.1045%
Variance	28.9736%^{2}
Skewness	-0.7719%
Kurtosis	4.5778%

The statistics are generated under the assumption that there is one IB with Rational Expectations. These are averages of 500 simulation rounds.

4.6 Conclusion

This paper uses Interactive Trial and Error Learning (Pradelski and Young, 2012) to model portfolio choice under Knightian Uncertainty. The model is calibrated against the real excess return of equity over bonds in the US financial market. It can match the mean and the variance of the excess return almost perfectly. It can also generate the threshold effect in the actual data at a sensible frequency.

4.7 Appendix

This appendix is based on Hansen (1997, 1999) and Franses and van Dijk (2000). Recall that

$$RER_t = (\phi_0^1+\sum_{j=1}^{12}\phi_j^1 RER_{t-j})\times 1(\underbrace{RER_{t-1}\leq\gamma}_{regime\,1})+(\phi_0^2+\sum_{j=1}^{12}\phi_j^2 RER_{t-j})\times 1(\underbrace{RER_{t-1}>\gamma}_{regime\,2})+e_t$$

is the model under consideration. Here, it will be shown how ϕ's and γ are estimated.

Let $\phi=(\phi_0^1,\phi_1^1,\dots,\phi_{12}^1,\phi_0^2,\phi_1^2,\dots,\phi_{12}^2)'$, $X_t=(1,RER_{t-1},RER_{t-2},\dots,RER_{t-12})$, and $X_t(\gamma)=(X_t\times 1(RER_{t-1}\leq\gamma)\;X_t\times 1(RER_{t-1}>\gamma))$. Then, the expression above can be rewritten as

$$RER_t = X_t(\gamma)\phi + e_t.$$

For a given value of γ, ϕ is estimated by the Ordinary Least Squares (OLS):

$$\hat{\phi}(\gamma)=\arg\min_{\phi}\sum_{t=1}^{T^e} e_t^2/T=(\sum_{t=1}^{T^e}X_t(\gamma)'X_t(\gamma))^{-1}(\sum_{t=1}^{T^e}X_t(\gamma)'RER_t)$$

where $e_t = RER_t - X_t(\gamma)\phi$ and T^e is the effective sample size ($299 - 12 = 287$ here). In order to estimate γ, it is necessary to determine a set of values from which it will be evaluated. This set, which is denoted by Γ, is constructed by (a) reordering $(RER_t)_{t=0}^{T^e-1}$ from the lowest to the highest value and (b) removing the bottom and the top $\alpha\%$ quantiles so that each regime contains at least $\alpha\%$ of observations. Here, $\alpha = 10\%$ following Hansen (1999). For a given $\hat{\phi}(\gamma)$, compute the sample variance of e_t which is $\hat{\sigma}^2(\gamma) = \sum_{t=1}^{T^e} \hat{e}_t^2 / T^e$ where $\hat{e}_t = RER_t - X_t(\gamma)\hat{\theta}(\gamma)$. Then,

$$\hat{\gamma} = \arg\min_{\gamma \in \Gamma} \hat{\sigma}(\gamma)$$

which is pinned down by grid search over Γ. Finally, $\hat{\phi} = \hat{\phi}(\hat{\gamma})$ is chosen for inference. $\hat{\phi}$ has the usual asymptotic properties of the OLS estimators.

4.8 References

1. Adam, K., Marcet, A., and Nicolini, J. P. (2011). "Stock market volatility and learning," CEP Discussion Papers dp1077, Centre for Economic Performance, LSE.

2. Amihud, Y. (2002). "Illiquidity and stock returns: cross-section and time-series effects," Journal of Financial Markets, Elsevier, vol. 5(1), pages 31-56, January.

3. Campbell, J. Y. (2003). "Consumption-based asset pricing," in Constantinides, G. M., Harris, M., and Stulz, R. M. (ed.), Handbook of the Economics of Finance, Elsevier, vol. 1, chapter 13, pages 803-887.

4. Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). "The econometrics of financial markets," Princeton Univeristy Press.

5. Franses, P. H., and Van Dijk, D. (2000). "Non-linear time series models in empirical finance," Cambridge University Press.

6. Hansen, B. E. (1997). "Inference in TAR models," Studies in Nonlinear Dynamics and Econometrics, Berkeley Electronic Press, vol. 2(1), pages 1-14, April.

7. _______ (1999). "Testing for linearity," Journal of Economic Surveys, Wiley Blackwell, vol. 13(5), pages 551-76, December.

8. Jarque, C. M., and Bera. A. K. (1987). "A test for normality of observations and regression residuals," International Statistical Review, International Statistical Institute, vol. 55(2), pages 163–172, August.

9. Knight, F. H. (1921). "Risk, uncertainty, and profit," Hart, Schaffner, and Marx.

10. Lippman, S. A. and McCall, J. J. (1986). "An operational measure of liquidity," American Economic Review, American Economic Association, vol. 76(1), pages 43-55.

11. Mehra, R. and Prescott, E. C. (1985). "The equity premium: A puzzle," Journal of Monetary Economics, Elsevier, vol. 15(2), pages 145-161, March.

12. _______ (2003). "The equity premium in retrospect," in Constantinides, G. M., Harris, M., and Stulz, R. M. (ed.), Handbook of the Economics of Finance, Elsevier, vol. 1, chapter 14, pages 889-938.

13. Newey, W. K. and West, K. D. (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," Econometrica, Econometric Society, vol. 55(3), pages 703-08, May.

14. Pradelski, B. S. and Young, H. P. (2012). "Learning efficient Nash equilibria in distributed systems," Games and Economic Behavior, Elsevier, vol. 75(2), pages 882-897, July.

15. Young, H. P. (2009). “Learning by trial and error,” Games and Economic Behavior, Elsevier, vol. 65(2), pages 626-643, March.

Chapter 5

Macroeconomic Dynamics under Intergenerational Forecasting Competition

Abstract

"This Time Is Different Syndrome" of Reinhart and Rogoff (2009) is explored in a setting where model agents are learning under Knightian Uncertainty. The agents are grouped into different generations and their models compete in terms of forecasting power. The predecessor's model is discarded together with the data set used to estimate it when its forecasting power is worse than the current generation's model. This loss of relevant and useful data is rooted in focusing only on forecasting well in the short-run. By shifting the weight towards finding the true model of the economy, this problem can be substantially reduced.

5.1 Introduction

In their historical explorations of economic crises, Reinhart and Rogoff (2009) conclude that

> "No matter how different the latest financial frenzy or crisis always appears, there are usually remarkable similarities with past experience from other countries and from history."

The essence of their argument is that economic crises have predictable patterns despite time and space differences. They claim that the failure to learn from history is at the heart of recurrent economic crises. This phenomenon is referred to as "This Time Is Different Syndrome" in their book. The key insight this paper takes away from their work is that the world is more stationary than what people perceive it to be and the failure to exploit this stationary can be an important mechanism by which economic shocks get reinforced. This paper explores this feature by introducing the agents whose model specification, evaluation, and selection activities can lead them to disregard the past data.

In this paper, the economy is modeled as a stationary stochastic process over time. "This Time Is Different Syndrome" takes the form of throwing away a data set that is relevant and useful for forming expectations about the future because a model agent falsely believes that the structure of the economy has changed in its time. It will be shown that this is rooted in focusing only on forecasting well in the short-run. The downside of focusing on forecasting power has been known since the time of Lucas (1976) who demonstrated this in the context of policy evaluation (the famous "Lucas Critique"). Sargent (1999) provides a clear illustration of this problem in relation to traditional Keynesian macroeconometric models. Recently, Edge and Gurkaynak (2011) showed that evaluating Dynamic Stochastic General Equilibrium models in terms of forecasting power is not fruitful. It will be shown that shifting the weight towards finding the true model of the economy can lessen "This Time Is Different Syndrome" in the context of the model in this paper.

The model agents in this paper face Knightian Uncertainty (Knight, 1921) in the sense that they do not know the true model of the economy. However, they are informed of a set of models which contains the true model of the economy. Thus, Knightian Uncertainty here is close to the one adopted by Robust Control Theory (see Hansen and Sargent, 2008) in which model agents deal with perturbations of a reference model. The nature of model uncertainty here is different from the one in the previous two chapters in which not even a set of models is provided to the agents.

The economy is occupied by finitely-lived agents which are classified into generations based on when they are born. Each generation comes up with a model that best describes its lifetime experience. This implies that different generations can believe in differently-specified models. These models compete over time in terms of their forecasting power. When the predecessor's model is shown to be forecasting worse than the current generation's model, it is thrown away together with the data set used to estimate it. This truncation of the data set is how "This Time Is Different Syndrome" manifests itself in this paper. This paper considers several model specification, evaluation, and selection methods which differ in their emphasis on forecasting well. It will be shown that focusing only on forecasting well in the short-run can lead to frequent rejections of the true model of the economy. The key result in this paper is that shifting the weight towards finding the true model of the economy, which is modeled by using Akaike and Schwarz Information Criterion (Akaike, 1974; Schwarz, 1978) for the purpose of model specification, can substantially reduce this problem. The modeling strategy in this paper is closely related to Econometric Learning Theory (see Evans and Honkapohja, 2009) in the sense that the model specification, evaluation, and selection methods here use the least squares method in order to take data to a model. A parallel between the modeling strategy here and econometric learning with a constant gain (see Barnett and Ellison, 2013) will be drawn later in this paper. Branch and Evans (2007), Cho and Kasas (2011), and Chung and Xiao (2013) are some examples where econometric learning is confounded by model selection. Chung and Xiao (2013) is particularly pertinent because model selection there is based on Schwarz Information

Criterion (Schwarz, 1978).

In contrast to the Rational Expectations setting in which model agents carry out optimizing and forecasting jointly, here they are separated into optimizers and forecasters. Whereas forecasters specialize in providing forecasts to optimizers and selecting models based on which these forecasts are generated, optimizers specializing in solving the optimal consumption, saving, and investment problem given these forecasts. As in the previous two chapters, the optimal investment problem takes the form of portfolio choice problem. The representative household in each generation is modeled as a team of the representative forecaster and the representative optimizer. The separation of optimizing and forecasting has a precedence given by Nerlove and Bessler (2001) who demonstrated the usefulness of this modeling strategy. For the purpose of tractability, this paper uses a simple setup in which returns on all assets are exogenous. Turning off feedback from the beliefs of the agents to the returns on assets allows a simple characterization of the Rational Expectations Equilibrium as the forecaster does not need to internalize the effect of its forecast on the portfolio choice by the optimizer which in turn influences its forecast and so on. Such is the case for the optimzer as well. The Rational Expectations Equilibrium is used as a benchmark throughout the paper. Overall, the model in this paper serves as an interesting platform upon which more realistic mechanisms can be built.

The paper is structured as follows: section 5.2 presents the main model; section 5.3 provides simulation results for the main model; section 5.4 extends the main model; section 5.5 concludes.

5.2 Model

5.2.1 Setup

In the economy, two assets are available for investment: safe (S) and uncertain (U). Let R_{t+1}^S denote the gross return on S in time period $t+1$ and R_{t+1}^U the gross return on U in time period $t+1$. As the names suggest, $R_{t+1}^S = R^S$ for all t whereas

$R_{t+1}^U = \mu(1-\rho) + \rho R_t^U + \varepsilon_{t+1}$ for all t where $|\rho| < 1$ and ε is an i.i.d. Gaussian white noise with the standard deviation of σ_ε. So, both assets evolve exogenously. For simplicity, R^S is normalized to 1. One may think of S as risk-free government bonds and U as equities.

The economy is occupied by finitely-lived households. The households are active for $T+1$ time periods and grouped into a generation based on when they are born. Each generation contains a continuum of households of unit measure. To make things more tractable, it is assumed that a new generation is born to replace an old generation only in the latter's last time period so that they overlap by only one period. This implies that there is only one generation to consider in each time period except in these last time periods.

For concreteness, let us work with a generation that is born in time period t and retires in time period $t+T$. Each household in this generation has a utility function $\sum_{s=0}^{T} \beta^s \left[\frac{c_{t+s}^{1-\sigma}-1}{1-\sigma}\right] + \beta^T \chi \left[\frac{(w_{t+T})^{1-\gamma}-1}{1-\gamma}\right]$ where c_{t+s} is a consumption in time period $t+s$, w_{t+T} is a bequest to the next generation (which is born in time period $t+T$) in time period $t+T$, $0 < \beta < 1$ is a discount factor, $\sigma > 0$ and $\gamma > 0$ are the coefficients of relative risk-aversion, and $\chi > 0$ is an intensity parameter for the bequest to the next generation. Each generation is modeled with a representative household.

In a typical setup, the representative household maximizes its utility function given the true models of R^S and R^U, i.e., its optimization problem is based on Rational Expectations[1]. Here, the representative household does not know how the true model of R^U is specified: it faces a specific form of Knightian Uncertainty[2] (Knight, 1921) which is rooted in the model uncertainty about R^U. The extent of the model uncertainty here is greater than the one under the usual econometric learning setup in which model agents are endowed with correct model specifications (but not correct model parameter values).[3]

Unlike the majority of macroeconomic models in which the representative household

---

[1] To be specific, it is a strong form of Rational Expectations a la Muth (1961).

[2] Chapter 3 provides a general discussion about Knightian Uncertainty.

[3] See Evans and Honkapohja (2009) for an up-to-date summary of the literature.

is a single agent that performs both forecasting and optimizing, this paper models the representative household as a team of forecasters and optimizers which specialize in the designated tasks. Nerlove and Bessler (2001) is an example in which a model agent's optimization task is carried out independently from its forecasting task. They refer to this as "separation of expectations and optimizing behavior" and demonstrate its usefulness based on a number of theoretical models. This modeling strategy is also supported by experimental evidence. Bao, Duffy, and Hommes (2013) ask whether economic agents can converge to a Rational Expectations Equilibrium (equilibrium in the sense that it is optimal) when the true model of the economy is not known to them and answer this question by carrying out experiments. They use a cobweb economy which is the most frequently used setup in the econometric learning literature. They consider four experimental treatments: (a) subjects forecast only (computer optimizes), (b) subjects optimize only, (c) subjects do both forecasting and optimizing, and (d) subjects form teams of two in which one does forecasting and the other optimizing. They show that subjects converge to a Rational Expectations Equilibrium under all four treatments. However, (c) has the slowest rate of convergence. Moreover, its form of convergence is weak in the sense that the estimated value of the Rational Expectations Equilibrium (based on the experimental data) is different from the correct value at 5% level of statistical significance. Such is not the case for (a) and (b) and is infrequently the case for (d). The outcomes under (a), (b), and (d) are in line with the old wisdom in economics that specialization is beneficial. In reference to the superiority of (d) over (c), the authors say "two heads are better than one in finding Rational Expectations Equilibrium."

The forecasters and the optimizers in the representative household are modeled by representative agents as well. The representative forecaster is referred to as "Analyst" and the representative optimizer as "Manager" henceforth. This taxonomy reflects tasks performed by financial analysts and fund managers in many investment banks. However, it is not necessary to interpret Analyst or Manager as a person. For instance, Manager can be thought of as an automated trading algorithm which accounts for substantial trading volumes in various financial markets today. In this case, the setup here is also consistent with experimental treatment (a) in the previous paragraph. Irrespective of how it is motivated, the key feature of the model in this paper is that a single representative household no long carries out both forecasting and optimizing.

Let $R^{U,F}$ stands for a forecast of R^U. In time period $t+s$, Analyst provides $\{R^{U,F}_{t+s+j}\}_{j=1}^{T-s}$ to Manager where $1 \leq s \leq T-1$. Given this, Manager solves the optimal consumption/saving/investment problem. In what follows, these are discussed in details.

5.2.2 Analyst

As mentioned in the previous subsection, Analyst can be thought of as a financial analyst working for a bank. There is much evidence that financial analysts are not like academic economists. For instance, financial analysts may be less interested in discovering the true model of the economy which is evidenced by the widespread use of technical analysis in financial industry. This can arise due to many reasons such as differences in incentives[4] or short-termism[5]. De Bondt and Thaler (1985; 1990) provide strong evidence that financial market participants such as security analysts tend to overreact. This is unlikely if they believe that their models provide accurate descriptions of financial markets. Adapting "cognitive consistency principle" which states that an economic model's agents are as good as economists (see Evans and Honkapohja, 2009), Analyst in this paper is modeled as a financial analyst whose primary objective is to forecast well in the short-run. One implication is that Analyst can do things that may be ad-hoc from the viewpoint of an academic economist.

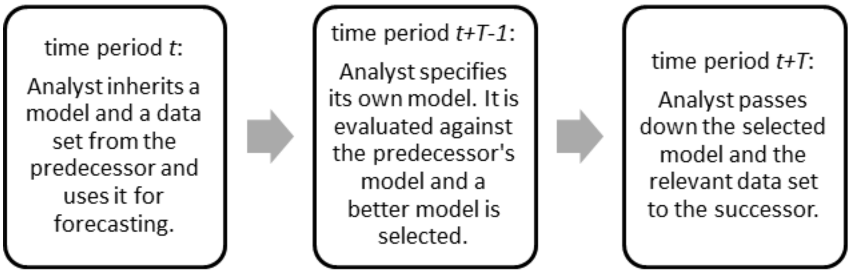
In what follows, let us consider Analyst of the generation that is born in time period t and retires in time period $t+T$. An overview of Analyst's life is provided first before going into the details. At the beginning of time period t, Analyst inherits a model of R^U as well as the data set used to estimate it from the previous generation's

---

[4] See Ramnath, Rock, and Shane (2008) for a survey of financial analyst forecasting literature which includes how incentives can influence forecast accuracy.

[5] See Rappaport (2005).

Figure 5.1: Analyst's Life



The figure summarizes the life of Analyst that is born in time period t and retires in time period $t + T$.

Analyst (so that it is possible to update the inherited model if necessary). This model is called the Predecessor's Model (PM henceforth). From time period t to $t + T - 1$, Analyst uses PM to generate forecasts of R^U or $R^{U,F}$s. These are used by Manager of the same generation for solving its optimal control problem. At the end of time period $t + T - 1$, which is just before its retirement, Analyst comes up with its Own Model (OM henceforth) of R^U. Because the true functional form of R^U is unknown, Analyst has to decide on a specification of OM first before estimating its parameters. The estimation involves only R^Us that have been directly observed by Analyst (as OM captures its own experience). OM is evaluated against PM and a better model is selected as the next generation's PM. It is passed down to the next generation's Analyst when the current generation's Analyst retires in time period $t + T$. If the current generation's PM is chosen as the next generation's PM, the data set that the next generation's Analyst receives includes the data set from the previous generation's Analyst as well. However, if OM is chosen as the next generation's PM, the data set consists of only R^Us that the current generation's Analyst has observed directly. Figure 5.1 summarizes Analyst's life.

Two motivations are given for the truncation of the data set that arises when OM

is chosen as the next generation's PM. First is Reinhart and Rogoff (2009)'s "This Time Is Different Syndrome." In their study, they collect a long time series data about various macroeconomic variables across many different countries. They show that economic crises have predictable patterns regardless of time and space differences and this suggests that the world is more stationary than commonly perceived. They argue that people fail to learn from history because they believe that the world they are living in is different from the past. This is what they refer to as This Time Is Different Syndrome. Because R^Us follow a stationary $AR(1)$ process, the truncation of the data set can be considered as an instance of This Time is Different Syndrome. Second, it is well-known that econometric learning with a constant gain parameter is equivalent to discounting the past data at a geometric rate in the least squares estimation (see Carceles-Poveda and Giannitsarou, 2007; Barnett and Ellison, 2013). Given that the truncation of the past data is a more extreme form of discounting the past data, Analyst in this paper can be thought of as engaged in econometric learning in which weights on data points in the least squares estimation evolve discontinuously. As usual, this can be interpreted as a reaction to possible non-stationarity. Note that the truncation of the data set here emerges endogenously through Analyst's model selection between OM and PM.

In what follows, Analyst's life will be described in detail.

1. Time period t to $t+T-1$: Analyst, who is born in time period t, inherits PM and the relevant data set from Analyst of the previous generation. PM is used to produce $R^{U,F}$s which are used by Manager of the same generation as an input to its optimal consumption/saving/investment problem. To be specific, Analyst provides $\{R^{U,F}_{t+s+j}\}_{j=1}^{T-s}$ to Manager in time period $t+s$. Whenever possible, the actual values of R^Us are used to generate $R^{U,F}$s, i.e., in time period $t+s$, the actual values of R^Us up to time period $t+s$ are used to generate $\{R^{U,F}_{t+s+j}\}_{j=1}^{T-s}$. For instance, for an estimated $AR(1)$ model, $R^{U,F}_{t+s+1} = \hat{a}_0 + \hat{a}_1 R^U_{t+s}$ and $R^{U,F}_{t+s+j} = \hat{a}_0 + \hat{a}_1 R^{U,F}_{t+s+j-1}$ for $j \geq 2$ where $\hat{a}$s are the estimates. However, Analyst uses PM without updating its estimates using R^Us from time period t to $t+T-1$

($\hat{a}$s in the example). Because any model manifests a belief about its object of interest, this lack of updating can be interpreted as a resistance to changes of beliefs about how R^Us are generated.[6] One obvious way to rationalize this setup is introducing a high updating cost which may be either physical or psychological in nature or both. This phase of Analyst's life corresponds to the first box in Figure 5.1.

2. At the end of time period $t+T-1$: Before its retirement in time period $t+T$, Analyst forms its own belief about how R^Us are generated by specifying its own model of R^Us which are based on its own observations of R^Us. As mentioned above, this is referred to as OM. Analyst evaluates PM and OM and selects one of them as a model that better characterizes its lifetime experience. The specification of OM, the evaluation of OM and PM, and the selection of the next generation's PM start by splitting the sample of R^Us from time period t to $t+T-1$ into two: $S_1 = \{R^U_{t+s}\}_{s=0}^{P}$ and $S_2 = \{R^U_{t+s}\}_{s=P+1}^{T-1}$ where $0 < P < T-1$. S_1 is used for estimation and S_2 for out-of-sample forecasting. Recall that the true model of R^Us is $AR(1)$ whose error terms are i.i.d. Gaussian white noises. In order to specify OM, Analyst considers a set of candidate models $M_K = \{AR(p) : 1 \leq p \leq K \text{ and error terms are i.i.d. Gaussian white noises}\}$ where p is a lag length and K is a positive integer. This set contains the true model of R^Us. Analyst estimates the models in M_K using S_1 and chooses the best model as OM (more on this below).[7] The evaluation and the selection are based on accuracy of out-of-sample forecasts in S_2. Mean Squared Prediction Error ($MSPE$) for S_2, which takes the form $\frac{\sum_{s=P+1}^{T-1}(R^U_{t+s}-R^{U,F}_{t+s})^2}{T-P-1}$ for a given $\{R^{U,F}_{t+s}\}_{s=P+1}^{T-1}$, is used to evaluate PM and OM. $MSPE$ of PM and OM are denoted by $MSPE^{PM}$ and $MSPE^{OM}$ respectively. The selection between PM and OM involves comparing $MSPE^{PM}$ and $MSPE^{OM}$ (more on this below). The selected

---

[6] See Armenakis and Bedeian (1999) and Piderit (2000) for literature reviews about resistance to organizational changes.

[7] If several estimated models emerge as the best models, Analyst chooses the one with the shortest p as OM. This becomes very relevant in section 5.4. See Granger (1999) for an argument in favor of parsimonious models.

model is passed down to the next generation's Analyst as its PM. Given that the current generation cares about the next generation (which is reflected in its utility function), it is reasonable that the model evaluation and selection are based on out-of-sample forecast accuracy as the next generation's Analyst will produce $R^{U,F}$s using what the current generation's Analyst passes down as its PM. As shown by Clark (2004), out-of-sample forecast comparisons can reduce overfitting as well and this is a valid concern here due to the form of M_K. The choice of $MSPE$ as a measure of out-of-sample forecast accuracy is due to its (a) optimality when the forecast errors are normally distributed[8] and (b) frequent use by professional forecasters. Note that when PM and OM have an identical specification, the model evaluation and selection reduces to an informal test of parameter drifts. This phase of Analyst's life corresponds to the second box in Figure 5.1.

3. At the beginning of time period $t + T$: Analyst re-estimates the selected model using the relevant data set. When PM from the previous generation is selected as the next generation's PM, it is re-estimated using the data set from the previous generation's Analyst and $S_1 \cup S_2 = \{R^U_{t+s}\}_{s=0}^{T-1}$ which has been observed by the current generation's Analyst directly. When OM is selected as the next generation's PM, it is re-estimated using only $S_1 \cup S_2$. The selected model is passed down to the next generation's Analyst as its PM together with the relevant data set. Analyst retires afterward. This phase of Analyst's life corresponds to the last box in Figure 5.1.

Following the usual practice in econometric learning literature (see Evans and Honkapohja, 2009), all the models are estimated by the least squares method. In what follows, several model specification, evaluation, and selection procedures will be presented. Broadly speaking, they can be grouped into two approaches:

1. Non-hypothesis Testing Approach:

---

[8] See Chen and Yang (2004).

- Specification and evaluation:

  Analyst estimates the models in M_K using S_1. Then, these estimated models are used to produce K batches of $\{R^{U,F}_{t+s}\}^{T-1}_{s=P+1}$. While the actual values of R^Us in S_2 are used as right-hand-side variables for this purpose, they are not used to update the parameter estimates of these models, i.e., $R^{U,F}$s are generated under "fixed" forecasting scheme. After computing $MSPE$ for S_2 for each batch of $\{R^{U,F}_{t+s}\}^{T-1}_{s=P+1}$, the estimated model with the minimum $MSPE$ is specified as OM.

  Analyst also uses PM to produce $\{R^{U,F}_{t+s}\}^{T-1}_{s=P+1}$ and uses this to compute $MSPE$ for S_2. As above, $\{R^{U,F}_{t+s}\}^{T-1}_{s=P+1}$ is generated under fixed forecasting scheme.

- Selection:

  Analyst selects PM if $MSPE^{PM} \leq MSPE^{OM}$ and OM otherwise.

2. Hypothesis Testing Approach based on Diebold-Mariano-West test (Diebold and Mariano, 1995; West, 1996)[9]:

   (a) $MSPE$

   - Specification and evaluation:

     It is identical to the one under Non-hypothesis Testing Approach.

   - Selection:

     Analyst uses Diebold-Mariano-West test of predictive accuracy (Diebold and Mariano, 1995; West, 1996) to select between PM and OM. Let e^{PM}_{t+s} and e^{OM}_{t+s} denote forecast errors of PM and OM in time period $t+s$ respectively and let $d_{t+s} = (e^{PM}_{t+s})^2 - (e^{OM}_{t+s})^2$. Because e's are assumed to be i.i.d., d's are i.i.d as well. Analyst conducts the one-sided hypothesis testing of $H_0 : E(d_{t+s}) = 0$ v.s. $H_1 : E(d_{t+s}) > 0$ where $E(d_{t+s})$ is the expected value of d_{t+s}. $E(d_{t+s})$ is estimated

---

[9] West (2006) provides a comprehensive and practitioner-oriented summary of forecast evaluation literature.

by $\bar{d} = MSPE^{PM} - MSPE^{OM} = \frac{\sum_{s=P+1}^{T-1} d_{t+s}}{T-P-1}$. The test statistic is $t_{DMW} = \frac{\bar{d}}{\hat{\sigma}_d}$ where $\hat{\sigma}_d = \sqrt{\frac{2\pi\hat{g}(0)}{T-P-1}}$ and $\hat{g}(0)$ is the spectral density of d_{t+s} at zero frequency. Because ds are assumed to be i.i.d., $2\pi\hat{g}(0) = \hat{\gamma}_d(0) = \frac{\sum_{s=P+1}^{T-1}(d_t-\bar{d})^2}{T-P-1}$ where $\hat{\gamma}_d(0)$ is the sample variance of d_{t+s}. t_{DMW} asymptotically follows the standard normal distribution. If H_0 cannot be rejected at α level of statistical significance, Analyst selects PM. Otherwise, it selects OM.

(b) Akaike Information Criterion (AIC) (Akaike, 1974)

- Specification and evaluation:

  Analyst estimates the models in M_K using S_1 and computes AICs of these estimated models. Recall that $AIC = 2(\frac{p+1}{P-p}) + \ln(\frac{\sum_{s=p+1}^{P}(R_{t+s}^U - \hat{R}_{t+s}^U)^2}{P-p})$ where $\hat{R}_{t+s}^U$ is the fitted value of R_{t+s}^U in time period $t+s$. The estimated model with the minimum AIC is specified as OM.

  Then, Analyst computes $MSPE$ for S_2 for PM and OM under fixed forecasting scheme (refer to specification and evaluation step under Non-hypothesis Testing Approach for the details).

- Selection:

  It is identical to the one under (a) of Hypothesis Testing Approach.

(c) Schwarz Information Criterion (SIC) (Schwarz, 1978)

- The model specification, evaluation, and selection are identical to (b) under Hypothesis Testing Approach except that SIC is used instead of AIC. Recall that $SIC = \ln(P-p)(\frac{p+1}{P-p}) + \ln(\frac{\sum_{s=p+1}^{P}(R_{t+s}^U - \hat{R}_{t+s}^U)^2}{P-p})$.

It is a well-known result that there is a trade-off between AIC and SIC: while AIC is efficient (in the sense of minimizing the Kullback-Leibler distance asymptotically), SIC is consistent (in the sense of selecting the true model almost-surely asymptotically) (See Diebold (2012) for further details). This is why both information criteria are considered here. Chung and Xiao (2013) provides an example of econometric learning in which model selection is based on SIC.

5.2.3 Manager

Recall that Manager belongs to the generation that is born in time period t and retires in time period $t+T$. At the beginning of time period t, Manager allocates the bequest from the previous generation between consumption and saving. The saving is deposited in an investment account which cannot be withdrawn until retirement in time period $t+T$ (so it covers the whole life). Upon retirement, the full value of the investment account is passed down to the next generation as a bequest.

The key implication of the irreversibility of the saving at the beginning of time period t is that the optimal consumption and saving problem can be separated from the optimal investment problem which takes the form of portfolio choice problem. Whereas the portfolio choice problem requires a new series of $R^{U,F}$s every time period for determining weights on assets S and U in an investment portfolio, the optimal consumption and saving problem is solved once and for all in time period t given $\{R^{U,F}_{t+s}\}_{s=1}^{T}$ from Analyst. In what follows, each problem will be presented in turn.

5.2.3.1 Optimal Consumption/Saving Problem

Manager solves

$$\max_{\{c_{t+s}\}_{s=0}^{T}, y_t} \sum_{s=0}^{T} \beta^s \left[\frac{c_{t+s}^{1-\sigma} - 1}{1-\sigma} \right] + \beta^T \chi \left[\frac{(w^F_{t+T})^{1-\gamma} - 1}{1-\gamma} \right]$$

$$\text{s.t.} \sum_{s=0}^{T} c_{t+s} + y_t \leq w_t, w_t \geq 0, w^F_{t+T} = \prod_{s=1}^{T} R^F_{t+s} y_t, c_{t+s} \geq 0 \text{ for } 0 \leq s \leq T, \text{ and } y_t \geq 0$$

where y_t is the value of the investment account in time period t (which is also the saving in time period t), w_t is the bequest received from the previous generation in time period t (which is also the initial wealth), w^F_{t+T} is the forecast of the bequest given to the next generation in time period $t+T$, R^F_{t+s} is the forecast of gross return on investment in time period $t+s$. $R^F_{t+s} = (R^{U,F}_{t+s} x_{t+s-1} + R^S(1 - x_{t+s-1})) = ((R^{U,F}_{t+s} - R^S) x_{t+s-1} + R^S)$ where x_{t+s-1} is the proportion of y_{t+s-1} invested in U in time period $t+s-1$. x_{t+s-1} is the solution of the portfolio choice problem which is discussed in the next subsubsection.

Note that w^F_{t+T} also depends on $\{x_{t+s}\}_{s=0}^{T-1}$ through its dependence on $\{R^F_{t+s}\}_{s=1}^{T}$.

Because the utility function is strictly increasing in $\{c_{t+s}\}_{s=0}^{T}$ and y_t, the budget constraint binds and the non-negativity constraints for c_{t+s}s and y_t do not bind. Substituting the budget constraint and w^F_{t+T} into the utility function, the utility maximization problem above becomes

$$\max_{\{c_{t+s}\}_{s=0}^{T}} \sum_{s=0}^{T} \beta^s \left[\frac{c_{t+s}^{1-\sigma}-1}{1-\sigma}\right] + \beta^T \chi \left[\frac{(\prod_{s=1}^{T} R^F_{t+s}(w_t - \sum_{s=0}^{T} c_{t+s})^{1-\gamma} - 1}{1-\gamma}\right].$$

The first order conditions are

$$c_{t+s}^{-\sigma} = \beta^{T-s}\chi(\prod_{s=1}^{T} R^F_{t+s})^{1-\gamma}(w_t - \sum_{s=0}^{T} c_{t+s})^{-\gamma} \tag{5.1}$$

for $s = 0, \ldots, T$. (5.1) implies that

$$c_{t+s} = (\beta^{\frac{1}{\sigma}})^s c_t. \tag{5.2}$$

for $s = 0, \ldots, T$. Substituting (5.2) into $\sum_{s=0}^{T} c_{t+s}$,

$$\sum_{s=0}^{T} c_{t+s} = \sum_{s=0}^{T} (\beta^{\frac{1}{\sigma}})^s c_t = \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}} c_t. \tag{5.3}$$

Substituting (5.3) into (5.1) for $s = 0$,

$$c_t^{-\sigma} = \beta^T \chi(\prod_{s=1}^{T} R^F_{t+s})^{1-\gamma}(w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}} c_t)^{-\gamma}. \tag{5.4}$$

(5.4) can be solved for c_t using numerical methods. c_{t+s} for $s = 1, \ldots, T$ can be obtained by substituting the solution of (5.4) into (5.2). Finally, the budget constraint and (5.3) implies that

$$y_t = w_t - \sum_{s=0}^{T} c_{t+s} = w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}} c_t. \tag{5.5}$$

y_t can be obtained by substituting the solution of (5.4) into (5.5).

The comparative statics based on the solutions above show that the effect of a unit increase in $\prod_{s=1}^{T} R_{t+s}^{F}$ on c_{t+s} for $s = 0, \ldots, T$ and y_t depends on what value γ takes (refer to Appendix 5.6 for a full list of comparative statics). When $\gamma = 1$, these are zero because the income effect and the substitution effect offset each other exactly. This can be seen easily from (5.4) because $\prod_{s=1}^{T} R_{t+s}^{F}$ drops out when $\gamma = 1$. When $\sigma = \gamma = 1$, the closed-form solution for (5.4) exists which is

$$c_t = \frac{1}{\beta^T \chi + \frac{1-\beta^{T+1}}{1-\beta}} w_t = \phi(\beta, T, \chi) w_t. \tag{5.6}$$

Substituting (5.6) into (5.2) implies that for $s = 0, \ldots, T$,

$$c_{t+s} = \beta^s \phi(\beta, T, \chi) w_t. \tag{5.7}$$

Finally, substituting (5.6) into (5.5),

$$y_t = w_t - \frac{1-\beta^{T+1}}{1-\beta} \phi(\beta, T, \chi) w_t = w_t - \psi(\beta, T, \chi) w_t = (1 - \psi(\beta, T, \chi)) w_t. \tag{5.8}$$

$0 < \psi < 1$ by construction. (5.8) implies that

$$\sum_{s=0}^{T} c_{t+s} = \psi(\beta, T, \chi) w_t. \tag{5.9}$$

It is worth re-emphasizing that (5.8) and (5.9) are constant fractions ψ and $1 - \psi$ of w_t which do not depend on $\{R_{t+s}^{U,F}\}_{s=1}^{T}$ from Analyst. This is because the income effect and the substitution effect cancel each other exactly when $\gamma = 1$ as pointed out above.

5.2.3.2 Portfolio Choice Problem

Recall that the representative household's utility function is strictly increasing in w_{t+T}. Because $w_{t+T} = \max\{y_{t+T}, 0\}$ by assumption, Manager aims to maximize $y_{t+T} = \prod_{s=1}^{T} R_{t+s} y_t$.

Again, $R^F_{t+j+1} = ((R^{U,F}_{t+j+1} - R^S)x_{t+j} + R^S)$. In time period $t+s$, Manager solves

$$\max_{\{x_{t+j}\}_{j=s}^{T-1}} \prod_{j=s}^{T-1}((R^{U,F}_{t+j+1} - R^S)x_{t+j} + R^S)y_{t+s}$$

$$\text{s.t. } x_{t+j} \in [\underline{x}, \overline{x}] \text{ for a given } \{R^{U,F}_{t+s+j}\}_{j=1}^{T-s} \text{ from Analyst}$$

where $0 < \underline{x} < \overline{x} < 1$. These parameter restrictions rule out 0 and 1 as solutions so Manager cannot divest fully from either asset. Analyst solves this optimization problem for $s = 0, \ldots, T-1$. Note that Manager deals with different vintages of $R^{U,F}$s over time because Analyst uses the actual values of R^Us as the right-hand-side variables to produce $R^{U,F}$s (refer to subsection 5.2.2).

Given any $\{R^{U,F}_{t+s+j}\}_{j=1}^{T-s}$ and $\{x_{t+j+1}\}_{j=s}^{T-2}$, it is optimal to choose $x_{t+s} = \arg\max_{x_{t+s}\in[\underline{x},\overline{x}]}((R^{U,F}_{t+s+1} - R^S)x_{t+s} + R^S)$. This implies that Manager needs only one-period-ahead forecast of R^U for making its investment decision. The optimal portfolio choice rule is

1. $R^{U,F}_{t+s+1} > R^S \Rightarrow x_{t+s} = \overline{x}$.
2. $R^{U,F}_{t+s+1} < R^S \Rightarrow x_{t+s} = \underline{x}$.
3. $R^{U,F}_{t+s+1} = R^S \Rightarrow x_{t+s}$ is any $x_{t+s} \in [\underline{x}, \overline{x}]$.

To make this rule determinate, it is assumed that $x_{t+s} = \underline{x}$ when 3 happens. This can be thought of as a reflection of the risk-aversion in the utility function above. Then, it is possible to recast the optimal portfolio choice problem above as a binary choice problem where $x_{t+s} \in \{\underline{x}, \overline{x}\}$ for $s = 0, \ldots, T-1$ without loss of generality. The resulting rule is

$$x_{t+s}(R^{U,F}_{t+s+1}, R^S) = \begin{cases} \overline{x} & \text{if } R^{U,F}_{t+s+1} > R^S \\ \underline{x} & \text{if } R^{U,F}_{t+s+1} \leq R^S \end{cases}. \tag{5.10}$$

5.3 Simulation

In this section, outcomes under different model specification, evaluation, and selection methods are compared. The attention is limited to the case where $\sigma = \gamma = 1$ so that

(5.4) has a closed-form solution (which is given by (5.6)). For concreteness, this section makes use of the generation that is born in time period t and retires in time period $t+T$.

Naturally, the Rational Expectations case is considered as a benchmark. This corresponds to a setting in which the solution to the optimal consumption/saving problem are (5.6), (5.7), and (5.8) and the solution to the portfolio choice problem is (5.10) where $R^{U,F}_{t+s+1}$ is based on the true model of R^Us which is $R^U_{t+s+1} = \mu(1-\rho) + \rho R^U_{t+s} + \varepsilon_{t+s+1}$ (refer to subsection 5.2.1). Then, $R^{U,F}_{t+s+1} = \mu(1-\rho) + \rho R^U_{t+s}$ in time period $t+s$.

Recall that $M_K = \{AR(p) : 1 \leq p \leq K \text{ and error terms are i.i.d. Gaussian}\}$ is the set of candidate models that Analyst considers for specifying OM where K is an integer that defines the upper limit on the lag length of the autoregressive models in M_K. In what follows, $K = 1$ and $K = 12$ are considered. $K = 1$ corresponds to the situation in which Analyst always works with the correct model specification but is concerned about potential parameter drifts. This can be thought of as econometric learning (see Evans and Honkapohja, 2009) with an informal test of parameter drifts. $K = 12$ confronts Analyst with the model specification issue as well. Table 5.1 presents parameter values that are common to these two cases. Recall that ψ, which is the fraction of the initial wealth in time period t (w_t) devoted to the consumption ($\sum_{s=0}^{T} c_{t+s}$), is a function of χ (refer to equations (5.6) to (5.9)). χ is set so that a considerable proportion of the initial wealth is saved into the investment account ($y_t = w_t - \sum_{s=0}^{T} c_{t+s}$) in time period t. All the hypothesis tests are conducted at 5% level of statistical significance. w_0, which is the initial wealth of the very first generation, is normalized to 1.

As explained in subsection 5.2.2, Analyst uses $S_1 = \{R^U_{t+s}\}_{s=0}^{P}$ for estimating the models in M_K and $S_2 = \{R^U_{t+s}\}_{s=P+1}^{T-1}$ for generating out-of-sample forecasts based on which the model evaluation and selection are carried out. In what follows, the Monte Carlo simulations are conducted for various values of T and P. In particular, $T = 101$ and $T = 202$ are considered so that $S_1 \cup S_2$ comes with 100 and 200 observations respectively. The reported results are averages of 1000 simulation rounds.

Table 5.1: Parameter Values

Parameter	Classification	Explanation	Value
R^S	Technology	Gross return on S	1
μ	Technology	Mean of R^U	1.02
ρ	Technology	Autoregressive parameter of R^U	.99
σ_ε	Technology	Standard deviation of residual term of R^U (ε)	.05
β	Preference	Time discount factor	.99
χ	Preference	Intensity parameter for bequest	1000
σ	Preference	Coefficient of relative risk aversion for consumption	1
γ	Preference	Coefficient of relative risk aversion for bequest	1
α	Hypothesis testing	Level of statistical significance for hypothesis testing	.05
G	Simulation	Number of generations	100

5.3.1 $K = 1$

For $K = 1$, M_1 contains only $AR(1)$ which is the correct model specification for R^U. This implies that the three model specification, evaluation, and selection methods under the Hypothesis Testing Approach are identical (refer to subsection 5.2.2 to confirm this). Table 5.2 presents simulation results for $T = 101$ and $T = 201$ with Ps that correspond to 25%, 50%, and 75% of the length of $S_1 \cup S_2 = \{R^U_{t+s}\}_{s=0}^{T-1}$. P determines the length of S_1 which is used for estimating OM. "Data Continuity" measures the percentage of the times the predecessor's data set of the current generation is combined with its own observations of R^U ($S_1 \cup S_2$) and passed down to the next generation as its predecessor's data set. This is equivalent to the percentage of the times PM is selected over OM. Because both PM and OM have the correct specification and R^Us are stationary, a reasonable model specification, evaluation, and selection method should produce a high value of data continuity. For both Non-hypothesis Testing Approach and Hypothesis Testing Approach, Data Continuity falls as P becomes larger. This is simply due to the law of large numbers: $MSPE$ becomes more accurate as P becomes smaller because more out-of-sample forecast errors are available for its computation (to the extent that it does not undermine the accuracy of the estimated parameters for OM). Data Continuity is higher under Hypothesis Testing Approach than under Non-hypothesis Testing Approach for the obvious reason that $MSPE^{PM}$ must be sufficiently larger than $MSPE^{OM}$ for OM to be selected under the former (whereas

Table 5.2: Simulation for $K = 1$

$T = 101$		$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing	Data Continuity (%)	85.75	62.43	46.85
	Deviation from w^{RE} (%)	.5575	.9625	1.0946
Hypothesis Testing	Data Continuity (%)	98.45	94.72	91.38
	Deviation from w^{RE} (%)	.1134	.1915	.2351
$T = 201$		$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing	Data Continuity (%)	85.67	64.21	49.55
	Deviation from w^{RE} (%)	.4509	.7612	.8508
Hypothesis Testing	Data Continuity (%)	98.52	95.31	92.66
	Deviation from w^{RE} (%)	.0613	.1435	.1701

Data Continuity measures the percentage of the times PM is selected over OM so that the predecessor's data set is combined with own observations of R^U and passed down to the next generation. Deviation from w^{RE} is the percentage difference between the wealth under Rational Expectations and the wealth under the chosen method in the last time period of simulation. The reported numbers are averages of 1000 simulation rounds.

$MSPE^{PM} > MSPE^{OM}$ is sufficient for the latter).

"Deviation from w^{RE}" measures the percentage difference between the wealth under Rational Expectations and the wealth under the chosen model specification, evaluation, and selection method in the last time period of simulation. This is used as an indicator of the extent of the deviation from the Rational Expectations case. As expected, the wealth under Rational Expectations dominates all others. Under both Non-hypothesis Testing Approach and Hypothesis Testing Approach, Deviation from w^{RE} increases as P becomes larger. The extent of the deviation is larger under Non-hypothesis Testing Approach. These are also due to the negative relationship between P and the accuracy of $MSPE$ mentioned in the previous paragraph. Note that Deviation from w^{RE} is not very large in all the cases. Why does this happen? (5.10) implies that as long as $R^{U,F}_{t+s+1}$s produced by different models are on the same side of R^S, they lead to the same x_{t+s} thus the same R_{t+s+1}. To the extent that $R^{U,F}$s under a model specification, evaluation, and selection method are on the same side of R^S as $R^{U,F}$s under Rational Expectations, the resulting wealth under that method is similar to the one under Rational Expectations. And this is the case here because M_1 allows only the correct model specification for R^U. More fundamentally, the observed phenomenon is rooted in modeling R^U as an exogenous process. Because this rules out feedback from endogenous variables such as x, wealth is inherently robust to different model specification, evaluation, and selection methods. This will be seen more clearly once the results for $K = 12$ are presented below.

5.3.2 $K = 12$

Because M_{12} contains more than one model, the model specification, evaluation, and selection methods under Hypothesis Testing Approach are no longer identical. In addition to Data Continuity, the percentage of the times the selected model has the correct specification is reported as well. This is labeled as "Correct Specification." "Both" is the percentage of the times both Correct Specification and Data Continuity hold.

Table 5.3 gives results from simulation. Under both Non-hypothesis Testing Approach and Hypothesis Testing Approach, Correct Specification and Data Continuity tend to fall jointly as P becomes larger whereas Deviation from w^{RE} increases as P becomes larger. As explained in subsection 5.3.1, these happen because $MSPE$ becomes more accurate as P decreases.

Correct Specification and Data Continuity are jointly 100% for AIC and SIC under Hypothesis Testing Approach for $P = 25$ and $T = 101$. This implies that the use of these information criteria for model specification can overcome the model specification uncertainty about R^U and deliver better results than those in its absence (given in Table 5.2 of the previous subsection). Deviation from w^{RE} in this case is smaller than the one without the model specification uncertainty about R^U as well. As explained in Diebold (2012), this is because full-sample procedures like AIC and SIC are more powerful than procedures based on out-of-sample forecasts as a model specification tool (in both finite and large samples). Given that AIC and SIC are mainly concerned with finding the true model specification rather than forecasting well, the model specification based on them can be interpreted as Analyst behaving more like an economist rather than a financial analyst at least for the purpose of specifying OM. Correct Specification and Data Continuity based on AIC and SIC are quite good for other values of P and T

as well.

As was the case in the previous subsection, Hypothesis Testing Approach performs better than Non-hypothesis Testing Approach for the same reason that the former sets a higher bar for selecting OM over PM than the latter.

Again, Deviation from w^{RE} is not very large in all the cases. In addition to the reasons provided in the previous subsection, this is because M_{12} consists of models that nest the true model specification for R^U. This brings about tendency for $R^{U,F}$s based on them to be on the same side of R^S as $R^{U,F}$s under Rational Expectations (as long as the size of S_1 is large enough to make their estimates accurate).

5.3.3 Discussion

It was shown in subsection 5.3.1 that the model specification for OM that focuses on forecasting well (i.e., the minimum $MSPE$) in the short-run can lead to the frequent loss of valid data even when there is no model specification uncertainty. As shown in subsection 5.3.2, introducing the model specification uncertainty makes this way of specifying OM prone to passing down misspecified models as well. Even though the model selection based on hypothesis testing can mitigate these problems to an extent, it alone is not sufficient. Indeed, focusing only on forecasting well in the short-run is the reason why these happen. By shifting the weight towards finding the true model specification for R^U (i.e., the use of AIC and SIC for specifying OM), which is what economists would care more about, these problems can be substantially reduced.

5.4 Extensions

In this section, two extensions to the model specification, evaluation, and selection methods in subsection 5.2.2 are provided. The first is a wealth-based method. Given that the representative household's utility is increasing in its final wealth (which is given to the next generation as a bequest), it is reasonable to consider a model specification, evaluation, and selection method that maximizes wealth. The second extension

Table 5.3: Simulation for $K = 12$

$T = 101$			$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing		Correct Specification (%)	61.87	47.63	34.11
		Data Continuity (%)	81.31	51.79	31.82
		Both (%)	51.08	26.35	12.61
		Deviation from w^{RE} (%)	.7021	1.3737	1.8118
Hypothesis Testing	MSPE	Correct Specification (%)	80.10	59.24	39.96
		Data Continuity (%)	98.19	93.94	88.97
		Both (%)	79.00	56.46	36.55
		Deviation from w^{RE} (%)	.1206	.2474	.4298
	AIC	Correct Specification (%)	100	88.66	80.66
		Data Continuity (%)	100	97.23	94.34
		Both (%)	100	86.44	76.39
		Deviation from w^{RE} (%)	.0242	.1276	.1864
	SIC	Correct Specification (%)	100	96.06	95.90
		Data Continuity (%)	100	95.38	91.98
		Both (%)	100	91.71	88.27
		Deviation from w^{RE} (%)	.0242	.1730	.2251
$T = 201$			$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing		Correct Specification (%)	61.04	47.63	34.02
		Data Continuity (%)	81.76	54.11	34.72
		Both (%)	50.68	27.79	13.67
		Deviation from w^{RE} (%)	.5744	1.0416	1.3164
Hypothesis Testing	MSPE	Correct Specification (%)	82.26	60.44	42.38
		Data Continuity (%)	98.37	94.73	90.86
		Both (%)	81.24	58.04	39.52
		Deviation from w^{RE} (%)	.0809	.1906	.2928
	AIC	Correct Specification (%)	96.51	88.33	82.97
		Data Continuity (%)	99.30	97.29	95.06
		Both (%)	95.89	86.05	79.06
		Deviation from w^{RE} (%)	.0406	.0926	.1394
	SIC	Correct Specification (%)	98.38	97.42	97.46
		Data Continuity (%)	98.75	95.66	92.99
		Both (%)	97.17	93.22	90.64
		Deviation from w^{RE} (%)	.0554	.1371	.1675

Correct Specification measures the percentage of the times the selected model has the correct specification. Both gives the percentage of the times both Correct Specification and Data Continuity hold. Refer to Table 5.2 for other labeling conventions. The reported numbers are averages of 1000 simulation rounds.

introduces information friction which takes the form of the loss of data set(s) from the previous generation(s) and shows which model specification, evaluation, and selection methods are robust to this information friction. The model parametrization is identical to section 5.3 and all the reported results are averages of 1000 simulation rounds.

5.4.1 Wealth-based Method

For concreteness, this subsection works with the generation that is born in time period t and retires in time period $t+T$. Now, Analyst operates with the aim of finding a model that would have maximized the representative household's wealth in time period $t+T-1$ which is the last time period in which it is actively engaged in its work. This is essentially comparing the counterfactuals with the benefits of hindsight: Analyst asks what could have been the representative household's wealth in time period $t+T-1$ if it had produced $R^{U,F}$s based on a particular model in order to find the best model. Because the representative household's utility function is strictly increasing in its final wealth, which is passed down to the next generation as a bequest, a model specification, evaluation, and selection method based on the counterfactual wealth maximization can be justified on the ground that it is aligned with the goal of maximizing the representative household's utility. Again, Non-hypothesis Testing Approach and Hypothesis Testing Approach are considered in turn.

1. Non-hypothesis Testing Approach:

    - Specification and evaluation:

      Analyst estimates the models in M_K using S_1. Then, these estimated models are used to produce K batches of $\{R^{U,F}_{t+s}\}^{T-1}_{s=R+1}$ under fixed forecasting scheme using S_2 (refer to the specification and evaluation step under Non-hypothesis Testing Approach in subsection 5.2.2). For each $\{R^{U,F}_{t+s}\}^{T-1}_{s=R+1}$, Analyst computes the counterfactual wealth in time period $t+T-1$, which is denoted by w^{CF}_{t+T-1}, using (5.10). Then, Analyst specifies the estimated model with the maximum w^{CF}_{t+T-1} as OM. The maximum w^{CF}_{t+T-1} is denoted by w^{OM}_{t+T-1}. When several models result to the maximum w^{CF}_{t+T-1}, which can happen fairly frequently due to the form of M_K (refer to the related discussions in section 5.3), the most parsimonious model is specified as OM.

Analyst also uses PM to produce $\{R^{U,F}_{t+s}\}^{T-1}_{s=R+1}$ which is also generated under fixed forecasting scheme. Based on this and (5.10), Analyst computes the counterfactual wealth in time period $t+T-1$ which coincides with w_{t+T-1}. This is because $R^{U,F}_{t+s+1}$ for the portfolio choice problem in time period $t+s$ coincides with $R^{U,F}_{t+s+1}$ here (refer to subsubsection 5.2.3.2).

- Selection:

  Analyst selects PM if $w_{t+T-1} \geq w^{OM}_{t+T-1}$ and OM otherwise.

2. Hypothesis Testing Approach:

   - Specification and evaluation:

     The specification step is identical to the one under Non-hypothesis Testing Approach in the same subsection.

     $MSPE^{PM}$ and $MSPE^{OM}$ are computed using fixed forecasting scheme (refer to the specification and evaluation step under Non-hypothesis Testing Approach in subsection 5.2.2).

   - Selection:

     It is identical to the one under (a) of Hypothesis Testing Approach in subsection 5.2.2.

5.4.1.1 $K = 1$

Table 5.4 provides simulation results for $K = 1$. Recall that Analyst always works with the correctly specified model of R^U in this case. Unlike the results in Table 5.2 (of subsection 5.3.1), Data Continuity increases as P becomes larger under Non-hypothesis Testing Approach. Why does this happen? The estimates of OM and PM become more similar as P increases (recall that P is the size of the sample used for estimation of OM). Therefore, they tend to generate $R^{U.F}$s that recommend the same investment

Table 5.4: Simulation for $K = 1$

$T = 101$		$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing	Data Continuity (%)	91.41	92.64	95.71
	Deviation from w^{RE} (%)	.1286	.1054	.0785
Hypothesis Testing	Data Continuity (%)	98.45	94.72	91.38
	Deviation from w^{RE} (%)	.1134	.1915	.2351
$T = 201$		$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing	Data Continuity (%)	87.26	89.20	93.93
	Deviation from w^{RE} (%)	.1099	.0930	.0604
Hypothesis Testing	Data Continuity (%)	98.52	95.31	92.66
	Deviation from w^{RE} (%)	.0613	.1435	.1701

The labeling conventions are identical to Table 5.2.

portfolios which lead to the same level of wealth in time period $t + T - 1$. This is why Data Continuity is increasing in P. This also explains why Non-hypothesis Testing Approach does better than Hypothesis Testing Approach for $P = 75$.

Overall, Non-hypothesis Testing Approach in this subsubsection does better than Non-hypothesis Testing Approach in subsection 5.3.1. By construction, Hypothesis Testing Approach here are identical to Hypothesis Testing Approach there because M_1 contains only one model.

5.4.1.2 $K = 12$

Table 5.5 gives simulation results for $K = 12$. Recall that Analyst faces the model specification uncertainty in this case. Under both Non-hypothesis Testing Approach and Hypothesis Testing Approach, Correct Specification has a non-monotonic relationship with P: it is U-shaped for the values of P considered. This reflects the tension between the reliability of (measures of) forecasts and the convergence of the estimates of OM and PM. The table also shows that Hypothesis Testing Approach works better than Non-hypothesis Testing Approach across different values of P and T.

Overall, Hypothesis Testing Approach in this subsubsection works quite well. Its performance is comparable to AIC and SIC under Hypothesis Testing Approach in subsection 5.3.2. This suggests that the specification of OM based on the maximization of the counterfactual wealth is a useful complement to the specification of OM based

Table 5.5: Simulation for $K = 12$

$T = 101$		$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing	Correct Specification (%)	46.57	39.47	41.70
	Data Continuity (%)	85.10	85.16	90.29
	Both (%)	40.22	34.31	38.45
	Deviation from w^{RE} (%)	.2601	.3027	.2399
Hypothesis Testing	Correct Specification (%)	96.50	95.46	96.71
	Data Continuity (%)	98.63	95.31	91.85
	Both (%)	95.22	91.06	88.94
	Deviation from w^{RE} (%)	.1097	.1908	.2410
$T = 201$		$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing	Correct Specification (%)	37.75	35.02	38.06
	Data Continuity (%)	76.57	78.15	86.23
	Both (%)	29.60	28.06	33.59
	Deviation from w^{RE} (%)	.3001	.2857	.2079
Hypothesis Testing	Correct Specification (%)	96.02	94.12	94.96
	Data Continuity (%)	98.83	96.13	93.26
	Both (%)	94.95	90.55	88.62
	Deviation from w^{RE} (%)	.0619	.1394	.1697

The labeling conventions are identical to Table 5.3.

on the information criteria.

5.4.2 Information Friction

How robust are the model specification, evaluation, and selection methods above to the availability of information? In order to answer this question, this subsection introduces information friction in which the data set(s) from the predecessor(s) does not get passed down due to communication problems across the generations. Up until here, it has been the case that Analyst always receives both PM and the data set used for estimating it. Now, the predecessor's data set fails to arrive with an i.i.d. probability λ even though Analyst continues to receive PM all the time. The undelivered data set is not recovered anymore. This is different from the Sticky Information a la Mankiw and Reis (2002) because the friction there only delays information from being incorporated into expectations formation.

The augmentation of the information friction implies that if PM is selected over OM under the information friction, Analyst cannot reestimate PM with an additional data set $S_1 \cup S_2$ before passing it down to the next generation's Analyst because it does not have access to the data set used for estimating PM. Thus, the future generations may continue using this unupdated PM for a while. It will be shown that this can reinforce "This Time Is Different Syndrome" in the context of the model in this paper.

In what follows, $\lambda = .1$ and $\lambda = .3$ are considered. Because the outcomes under $T = 101$ and $T = 201$ are similar in the absence of the information friction, only the former will be pursued to save on computational time.

5.4.2.1 $K = 1$

Table 5.6 reports simulation results for both $\lambda = .1$ and $\lambda = .3$. Recall that there is no model specification uncertainty here. As expected, the information friction substantially worsens the performance of both $MSPE$-based and wealth-based model specification, evaluation, and selection method (whose details are provided in subsections 5.2.2 and 5.4.1 respectively). For Non-hypothesis Testing Approach, while the $MSPE$-based method has higher Data Continuity, the wealth-based method comes with lower Deviation from w^{RE}. Because M_1 contains only one model, the $MSPE$-based method and the wealth-based method are identical under Hypothesis Testing Approach.

5.4.2.2 $K = 12$

The simulation results in the presence of both the model specification uncertainty and the information friction are provided in Table 5.7 and Table 5.8. There are many things going on on these tables. Let us start by noting that AIC-based, SIC-based, and wealth-based model specification, evaluation, and selection method under Hypothesis Testing Approach are very good at detecting the correct model specification for R^U under both $\lambda = .1$ and $\lambda = .3$. Data Continuity decreases in all the cases given the nature of the information friction here. Overall, the methods based on the information criteria perform the best, especially in terms of Deviation from w^{RE}. The performance of the wealth-based method comes close to the performance of these methods.

Table 5.6: Simulation for $K = 1$

$\lambda = .1$			$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing	MSPE	Data Continuity (%)	35.07	38.75	33.28
		Deviation from w^{RE} (%)	.8434	1.0995	1.1808
	Wealth	Data Continuity (%)	27.99	26.18	20.46
		Deviation from w^{RE} (%)	.1949	.1520	.1290
Hypothesis Testing	MSPE	Data Continuity (%)	13.43	22.37	27.98
		Deviation from w^{RE} (%)	.2446	.3212	.3639
	Wealth	Data Continuity (%)	13.43	22.37	27.98
		Deviation from w^{RE} (%)	.2446	.3212	.3639
$\lambda = .3$			$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing	MSPE	Data Continuity (%)	25.07	29.58	25.87
		Deviation from w^{RE} (%)	1.0311	1.1954	1.2063
	Wealth	Data Continuity (%)	16.09	14.81	10.11
		Deviation from w^{RE} (%)	.2881	.2326	.2368
Hypothesis Testing	MSPE	Data Continuity (%)	7.86	14.80	18.80
		Deviation from w^{RE} (%)	.4909	.4457	.4704
	Wealth	Data Continuity (%)	7.86	14.80	18.80
		Deviation from w^{RE} (%)	.4909	.4457	.4704

MSPE and Wealth correspond to model specification, evaluation, and selection methods in subsections 5.2.2 and 5.4.1 respectively. Other labeling conventions are identical to Table 5.2.

The results above suggest that shifting the weight towards finding the true model specification for R^U, which is carried out by using the information criteria for specifying OM, can lessen the damage from the information friction as well.

5.5 Conclusion

This paper shows that focusing only on forecasting well in the short-run can lead to "This Time Is Different Syndrome" in which the relevant data are falsely discarded. This also increases the frequency with which wrong model specifications are selected. By shifting the weight towards finding the true model, which is modeled by using Akaike and Schwarz Information Criterion (Akaike, 1974; Schwarz, 1978) for the purpose of model specification, these problems can be substantially reduced.

The model in this paper provides an interesting platform upon which more realistic features can be built. A particularly interesting extension would be allowing the returns on some assets to be endogenous so that there is two-way feedback between the beliefs

Table 5.7: Simulation for $K = 12$ and $\lambda = .1$

			$P = 25$	$P = 50$	$P = 75$
Non-hypothesis Testing	MSPE	Correct Specification (%)	61.10	47.75	34.25
		Data Continuity (%)	37.09	35.06	24.11
		Both (%)	23.18	17.79	9.59
		Deviation from w^{RE} (%)	.9759	1.4893	1.8774
	Wealth	Correct Specification (%)	47.30	40.71	43.57
		Data Continuity (%)	34.20	34.16	29.40
		Both (%)	17.09	14.90	13.73
		Deviation from w^{RE} (%)	.3354	.4001	.3475
Hypothesis Testing	MSPE	Correct Specification (%)	78.70	59.15	42.14
		Data Continuity (%)	14.00	23.15	30.74
		Both (%)	10.69	13.48	13.21
		Deviation from w^{RE} (%)	.2567	.4166	.6483
	AIC	Correct Specification (%)	100	86.94	79.54
		Data Continuity (%)	7.00	16.74	23.16
		Both (%)	7.00	14.28	18.41
		Deviation from w^{RE} (%)	.0589	.2669	.3472
	SIC	Correct Specification (%)	100	95.74	95.34
		Data Continuity (%)	7.00	21.18	27.23
		Both (%)	7.00	20.21	26.01
		Deviation from w^{RE} (%)	.0589	.3147	.3587
	Wealth	Correct Specification (%)	95.22	94.49	95.99
		Data Continuity (%)	12.70	21.04	27.32
		Both (%)	12.07	19.87	26.09
		Deviation from w^{RE} (%)	.2485	.3459	.3866

The labeling conventions can be found in Table 5.3 and Table 5.6.

Table 5.8: Simulation for $K = 12$ and $\lambda = .3$

			$P = 50$	$P = 100$	$P = 150$
Non-hypothesis Testing	MSPE	Correct Specification (%)	61.16	47.94	34.30
		Data Continuity (%)	27.64	27.04	19.10
		Both (%)	16.38	13.55	7.25
		Deviation from w^{RE} (%)	1.1750	1.5964	1.9291
	Wealth	Correct Specification (%)	47.69	41.25	44.22
		Data Continuity (%)	22.18	21.98	17.44
		Both (%)	9.86	8.24	6.43
		Deviation from w^{RE} (%)	.4252	.4628	.4307
Hypothesis Testing	MSPE	Correct Specification (%)	67.33	56.16	39.43
		Data Continuity (%)	8.66	15.64	21.04
		Both (%)	5.34	8.00	7.27
		Deviation from w^{RE} (%)	.5104	.5507	.7565
	AIC	Correct Specification (%)	96.90	81.91	75.87
		Data Continuity (%)	.11	10.37	14.97
		Both (%)	0	8.03	11.06
		Deviation from w^{RE} (%)	1.3391	.4459	.4948
	SIC	Correct Specification (%)	96.90	93.88	94.87
		Data Continuity (%)	.11	13.68	18.07
		Both (%)	0	12.73	17.10
		Deviation from w^{RE} (%)	1.3391	.4309	.4528
	Wealth	Correct Specification (%)	90.94	93.40	95.59
		Data Continuity (%)	7.13	13.71	18.20
		Both (%)	6.42	12.69	17.24
		Deviation from w^{RE} (%)	.5250	.4874	.4899

The labeling conventions can be found in Table 5.3 and Table 5.6.

of the agents and the returns on these assets. This will be addressed in the future research.

5.6 Appendix

The comparative statics here are based on the solutions to the optimal consumption/saving problem in subsubsection 5.2.3.1. For each type of comparative statics, the first entry is based on (5.4), the second based on (5.5), and the third based on (5.2).

1. Comparative statics for w_t:

   (a) $\frac{\partial c_t}{\partial w_t} = \frac{\beta^T \chi(\prod_{s=1}^T R_{t+s}^F)^{1-\gamma}\gamma(w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}c_t)^{-\gamma-1}}{\sigma c_t^{-\sigma-1} + \beta^T \chi(\prod_{s=1}^T R_{t+s}^F)^{1-\gamma}\gamma(w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}c_t)^{-\gamma-1}(\frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}})} > 0$ (but less than 1 because $\frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}} > 1$).

   (b) $\frac{\partial y_t}{\partial w_t} = 1 - \underbrace{\underbrace{\frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}}_{>1}\underbrace{\frac{\partial c_t}{\partial w_t}}_{<1}}_{?} =?.$

   (c) $\frac{\partial c_{t+s}}{\partial w_t} = \underbrace{(\beta^{\frac{1}{\sigma}})^s}_{>0}\underbrace{\frac{\partial c_t}{\partial w_t}}_{>0} > 0$ for $s = 1, \ldots, T$.

2. Comparative statics for $\prod_{s=1}^T R_{t+s}^F$:

   (a) $\frac{\partial c_t}{\partial \prod_{s=1}^T R_{t+s}^F} = (\gamma-1) \times \frac{\beta^T \chi(\prod_{s=1}^T R_{t+s}^F)^{-\gamma}(w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}c_t)^{-\gamma}}{\underbrace{\sigma c_t^{-\sigma-1} + \beta^T \chi(\prod_{s=1}^T R_{t+s}^F)^{1-\gamma}\gamma(w_t - \frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}c_t)^{-\gamma-1}(\frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}})}_{>0}}.$

   (b) $\frac{\partial y_t}{\partial \prod_{s=1}^T R_{t+s}^F} = -\underbrace{\frac{1-(\beta^{\frac{1}{\sigma}})^{T+1}}{1-\beta^{\frac{1}{\sigma}}}}_{>0}\frac{\partial c_t}{\partial \prod_{s=1}^T R_{t+s}^F}.$

   (c) $\frac{\partial c_{t+s}}{\partial \prod_{s=1}^T R_{t+s}^F} = \underbrace{(\beta^{\frac{1}{\sigma}})^s}_{>0}\frac{\partial c_t}{\partial \prod_{s=1}^T R_{t+s}^F}$ for $s = 1, \ldots, T$.

   If $\gamma = 1$, $\frac{\partial c_t}{\partial \prod_{s=1}^T R_{t+s}^F} = 0 \Rightarrow \frac{\partial c_{t+s}}{\partial \prod_{s=1}^T R_{t+s}^F} = 0$ for $s = 1, \ldots, T$ and $\frac{\partial y_t}{\partial \prod_{s=1}^T R_{t+s}^F} = 0$.

   If $\gamma > 1$, $\frac{\partial c_t}{\partial \prod_{s=1}^T R_{t+s}^F} > 0 \Rightarrow \frac{\partial c_{t+s}}{\partial \prod_{s=1}^T R_{t+s}^F} > 0$ for $s = 1, \ldots, T$ and $\frac{\partial y_t}{\partial \prod_{s=1}^T R_{t+s}^F} < 0$.

$$\text{If } \gamma < 1, \ \frac{\partial c_t}{\partial \prod_{s=1}^{T} R_{t+s}^F} < 0 \Rightarrow \frac{\partial c_{t+s}}{\partial \prod_{s=1}^{T} R_{t+s}^F} < 0 \text{ for } s = 1, \ldots, T \text{ and } \frac{\partial y_t}{\partial \prod_{s=1}^{T} R_{t+s}^F} > 0.$$

As seen from 1, the effect of a unit increase in w_t on y_t is ambiguous because the direct positive effect of w_t on y_t is offset by the indirect negative effect of w_t on y_t through c_{t+s} for $s = 1, \ldots, T$. 2 shows that the effect of a unit increase in $\prod_{s=1}^{T} R_{t+s}^F$ on c_{t+s} for $s = 0, \ldots, T$ and y_t depends on what value γ takes. When $\gamma = 1$, these are zero because the income effect and the substitution effect offset each other exactly.

5.7 References

1. Akaike, H. (1974). “A new look at the statistical model identification,” Automatic Control, IEEE Transactions on, vol. 19(6), pages 716-723.

2. Armenakis, A. A. and Bedeian, A. G. (1999). “Organizational change: a review of theory and research in the 1990s,” Journal of management, vol. 25(3), pages 293-315.

3. Bao, T., Duffy, J., and Hommes, C. (2013). “Learning, forecasting and optimizing: an experimental study,” European Economic Review, Elsevier, vol. 61(C), pages 186-204.

4. Barnett, A. and Ellison, M. (2013). “Learning by disinflating,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 45(4), pages 731-746, June.

5. Branch, W. and Evans, G. W. (2007). “Model uncertainty and endogenous volatility,” Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 10(2), pages 207-237, April.

6. Carceles-Poveda, E. and Giannitsarou, C. (2007). “Adaptive learning in practice,” Journal of Economic Dynamics and Control, Elsevier, vol. 31(8), pages 2659-2697, August.

7. Chen, Z. and Yang, Y. (2004). “Assessing forecast accuracy measures,” Working Paper.

8. Cho, I. K. and Kasas, K. (2011). “Learning and model validation,” 2011 Meeting Papers 1086, Society for Economic Dynamics.

9. Chung, H. and Xiao, W. (2013). “Cognitive consistency, signal extraction, and macroeconomic persistence,” Working Paper.

10. Clark, T. E. (2004). “Can out-of-sample forecast comparisons help prevent overfitting?,” Journal of Forecasting, John Wiley & Sons, Ltd., vol. 23(2), pages 115-139.

11. De Bondt, W. F. M. and Thaler, R. (1985). “Does the stock market overreact?,” Journal of Finance, American Finance Association, vol. 40(3), pages 793-805, July.

12. _______ (1990). “Do security analysts overreact?,” American Economic Review, American Economic Association, vol. 80(2), pages 52-57, May.

13. Diebold, F. X. (2012). “Comparing predictive accuracy, twenty years later: a personal perspective on the use and abuse of Diebold-Mariano tests,” Working Paper 18391, National Bureau of Economic Research, Inc.

14. Diebold, F. X. and Mariano, R. S. (1995). “Comparing predictive accuracy,” Journal of Business & Economic Statistics, American Statistical Association, vol. 13(3), pages 253-63, July.

15. Edge, R. M. and Gurkaynak, R. S. (2011). “How useful are estimated DSGE model forecasts for central bankers?,” Brookings Papers on Economic Activity, Fall 2010, pages 209-244.

16. Evans, G. W. and Honkapohja S. (2009). “Learning and macroeconomics,” Annual Review of Economics, Annual Reviews, vol. 1(1), pages 421-451, June.

17. Granger, C. W. (1999). “Empirical modeling in economics: specification and evaluation,” Cambridge University Press.

18. Hansen, L. P. and Sargent, T. J. (2008). “Robustness,” Princeton University Press.

19. Knight, F. H. (1921). “Risk, uncertainty, and profit,” Hart, Schaffner, and Marx.

20. Lucas, R. (1976). “Econometric policy evaluation: a critique,” Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 1(1), pages 19-46, January.

21. Mankiw, N. G. and Reis, R. (2002). “Sticky information versus sticky prices: a proposal to replace the new Keynesian Phillips curve,” The Quarterly Journal of Economics, MIT Press, vol. 117(4), pages 1295-1328, November.

22. Muth, J. F. (1961). “Rational expectations and the theory of price movements,” Econometrica, Econometric Society, vol. 29(3), pages 315-335.

23. Nerlove, M. and Bessler, D. A. (2001). “Expectations, information and dynamics,” in Gardner B. L. and Rausser, G. C. (ed.), Handbook of Agricultural Economics, Elsevier, vol. 1, chapter 3, pages 155-206.

24. Piderit, S. K. (2000). “Rethinking resistance and recognizing ambivalence: a multidimensional view of attitudes toward an organizational change,” Academy of Management Review, vol. 25(4), pages 783-794.

25. Ramnath, S., Rock, S., and Shane, P. (2008). “The financial analyst forecasting literature: a taxonomy with suggestions for further research,” International Journal of Forecasting, vol. 24(1), pages 34-75.

26. Rappaport, A. (2005). “The economics of short-term performance obsession,” Financial Analysts Journal, vol. 61(3), pages 65-79.

27. Reinhart, C. M. and Rogoff, K. S. (2009). “This time is different: eight centuries of financial folly,” Princeton University Press.

28. Sargent, T. J. (2002). “The conquest of American inflation,” Princeton University Press.

29. Schwarz, G. (1978). “Estimating the dimension of a model,” The Annals of Statistics, vol. 6(2), pages 461-464.

30. West, K. D. (1996). “Asymptotic inference about predictive ability,” Econometrica, Econometric Society, vol. 64(5), pages 1067-84, September.

31. _______ (2006). “Forecast evaluation,” in Elliott, G., Granger, C., and Timmermann, A. (ed.), Handbook of Economic Forecasting, Elsevier, vol. 1, pages 99-134.