10-1-2014

Committee Decision Making in Monetary Policy

Hyungju Cha
Graduate Center, City University of New York

How does access to this work benefit you? Let us know!
Follow this and additional works at: http://academicworks.cuny.edu/gc_etds
Part of the Economic Theory Commons, and the Political Science Commons

Recommended Citation
Cha, Hyungju, "Committee Decision Making in Monetary Policy" (2014). CUNY Academic Works.
http://academicworks.cuny.edu/gc_etds/410

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Graduate Works by Year: Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
COMMITTEE DECISION MAKING IN MONETARY POLICY

By

HYUNJU CHA

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014
This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Merih Uctum

__________________________

Date                  Chair of Examining Committee

Merih Uctum

__________________________

Date                  Executive Officer

Professor   Merih Uctum

Professor   Thom Thurston

Professor   George Vachadze

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

Committee Decision Making in Monetary Policy

by

Hyungju Cha

Advisor: Professor Merih Uctum

One reason of the criticism about Taylor rule would be that the rule could not provide a micro foundation about the decision-making mechanism. because it is simply describing the policy process in terms of the short-term nominal interest rate. In order to reconcile the tension between critics and supporters of Taylor rule, chapter I employs the theory of team, which investigates how members in a team interact to determine optimal decisions. Using team theory, it is shown how optimal monetary policy is results from committee members' interactions.

In chapter II, if the government cannot recognizes that the monetary policy can affect the economy with a lag or if observed data for decision-making is real-time data, following Orphanides (2001, 2003), this policy is not effective as much as their expectation. From Friedman's metaphor of "Fool in the shower", it is reasonable to include time lags in committee decision-making. In chapter II, including lags in the information structure, recognition and administrative lags, optimal inertial monetary policy is derived.

In chapter III, FIML and SUR are employed to check if the actual data support the optimal monetary policy obtained under committee decision-making. The empirical evidence supports the inclusion of the lagged interest rate and the persistent shock in errors.
Acknowledgments

At the beginning of studying economics, there are many persons to support me. First of all, Jesus Christ is hidden supporter to finish my dissertation. I really appreciate his unbelievable plan for me. Second of all, without my family, especially, my mom and wife, I cannot finish the dissertation. I give the special gratitude to them for their countless help. Third of all, I thank Professors. All knowledge I know are from them. Fourth of all, whenever I lose my faith and direction, pastors help me correct my direction and faith. They are the compass in my life. I also thank pastors.

Finally, my new baby, Luke S. Cha, gives me the strength at the end of writing the dissertation. I really welcome to my new baby.
# Contents

Chapter 1. Committee Decision Making in Monetary Policy ................................. 1

1.1 Introduction ........................................................................................................ 1

1.2 Decision Making in the Committee ................................................................. 4

1.2.1 Theory of team ............................................................................................. 6

1.2.2 Communication pattern in a team ............................................................... 10

1.2.3 The optimal monetary policy under team situations .................................. 14

1.2.4 Solution Technique ................................................................................... 16

1.2.5 Numerical Solution to Derive the Optimal monetary Policy ..................... 22

1.3 Conclusion - The optimal monetary policy under team situation .................. 27

Chapter 2. Fool in the shower ................................................................................... 29

2.1 Introduction ....................................................................................................... 29

2.2 Numerical Solution to Derive the Optimal monetary Policy Under Lags ........ 33

2.3 Conclusion - Gradualism vs. Persistent shock ............................................... 39

Chapter 3. Numerical Estimation of the Gradualism ............................................. 40

3.1 Introduction ...................................................................................................... 40

3.2 Basic Model .................................................................................................... 42
3.3 Estimating the System.......................................................................................................... 45
3.4 Estimation Result................................................................................................................. 47
3.5 Estimating the System without the Persistent Shock........................................................... 49
3.6 Estimating the System without the Lagged Interest Rate.................................................... 51
3.7 Comparison with the Seemingly Unrelated Regression Estimation................................. 53
3.8 Business Cycle Sensitivity in the Policy............................................................................. 53
3.9 Conclusion: Empirical Approach........................................................................................ 56

Bibliography......................................................................................................................... 58
List of Tables

3.1 Quasi-FIML Estimation Results .......................................................................................... 48
3.2 Quasi-FIML Estimation Results without the Persistent Shock ........................................... 51
3.3 Quasi-FIML Estimation Results without the Lagged Interest Rate .................................... 52
3.4 SUR Estimation Results ...................................................................................................... 53
3.5 SUR Estimation Results without the Persistent Shock ....................................................... 54
3.6 SUR Estimation Results without the Lagged Interest Rate ............................................... 55
3.7 Quasi-FIML Estimation Results with the Sub-Sample above 6 Percent ......................... 56
3.8 Quasi-FIML Estimation Results with the Sub-Sample below 6 Percent ......................... 57
List of figures

1.1 Derivation of actual funds rate from estimated value from Taylor rule (Rudebusch, 2006)........... 4
1.2 An example of communication patterns (Bavelas, 1951).............................................................. 13
1.3 The procedure to find out the optimal solution (Chu, 1972)........................................................ 13
1.4 The communication pattern in committee decision-making under artificial FOMC.................. 15
1.5 Equivalent Controls (Chu, 1972).................................................................................................. 20

2.1 Decision-making procedure in artificial FOMC............................................................................. 32
2.2 Movement of Effective Federal Funds Rates................................................................................. 33

3.1 The model's fitting to the data........................................................................................................ 49
3.2 Movement of Unemployment Rates............................................................................................. 56
Chapter 1

Committee Decision Making in
Monetary Policy

1.1 Introduction

After Taylor (1993) suggested a simple framework describing the evolution of monetary policy in the U.S., this has had considerable influence on monetary economics. This is not only because Taylor was successful in explaining the monetary policy within Paul Volcker and Alan Greenspan era but also because this simple rule can describe policy process in terms of the short-term nominal interest rate. Even with the success in explaining the Federal Funds rate within Paul Volcker and Alan Greenspan era, there are two main shortcoming in Taylor-type rule. One is theoretical problem, and the other is empirical one.

From theoretical point of view, Taylor rule is simply describing the decision-making process only in terms of short-term nominal interest rates. However, it is too simple to capture the property of the actual decision-making process. Of course, this can be effective framework. This is because Taylor rule can be derived from minimizing the social welfare loss in which it has some components.
such as output gap and inflation gap. Due to this reason, Taylor rule can be seen as having micro-foundation. As known, the social planner problem is based on Tinbergen’s principal idea. (See Tibergen (1952)), that is, the government’s objective or welfare function could be written as endogenous variable like inflation, unemployment, or output gap if every economic system is complete. Therefore, this approach is very simple and attractive. But, as pointed out by Theil (1956), this approach has some problems. One of them is this approach assumes that policy-makers’ utility is satisfied by corresponding to target variables. The problem implies that this policy function is ad-hoc. In contrast with this, Woodford (2003) suggest micro-founded loss function or utility-based loss function which includes second-order approximation of agent’s utility maximization. However, since this approach is still minimizing the social welfare loss, it is still subject to Theil’s criticism. Cochrane (2011) argues that without assuming the orthogonality between policy instrument variables and the disturbance, it could not be available to measure the policy response, for instance, coefficients on output gap and inflation gap. Since output and inflation could be related to each other through the Phillips curve, both variables are affected by same disturbance. Therefore, the correlation of instruments in Taylor rule with errors could not be removed. This is known as identification problem. Therefore, price determinancy is not available due to identification problem, that is, there are multiple equilibria. This is because the condition for a unique stationary equilibrium depends on coefficients in Taylor rule. If policy disturbance is forecasting error, however, determinancy and identification could be possible since policy instruments are orthogonal to forecasting error. The problem in here is how policy disturbance could be forecasting error.\footnote{In addition to this, Benhabib, Schmitt-Grohe and Uribe (2001) pointed out that Taylor rule leads to global indeterminacy even though equilibrium is locally unique.} According to Rudebusch (2006), moreover, Taylor rule provides incomplete description of monetary policy. It provides a basic narrative history of how the Federal Reserve responses to macroeconomic...
development factors such as financial market factors in addition to current output gap and inflation. (These macroeconomic developments appear in figure 1.1.) In figure 1.1, without including financial frictions in Taylor rule, both inertial Taylor rule and non-inertial one are deviated from actual Federal Funds rate when there are some financial frictions. Therefore, only output gap and inflation in the rule is not enough to describe the movement of Federal Funds rate and it is needed to include macroeconomic development in the rule without amending the Fed’s objective function. In this vein, it would be reasonable to include additional or different variables in Taylor rule.

In empirical point of view, Orphanides (2001) argues that simple policy rule is based on the timeliness of data availability and ignore the difference between initial data and subsequent data revision. In addition, Orphanides (2003) also argue that noisy information regarding the measurement of economic activity for the evaluation of monetary policy can not account for the actual economic activity for the evaluation. Therefore, although the rule may appear to capture the movement in actual data, it can not provide actual economic situation if the revised data are employed. In this vein, the success of Taylor rule to explain the movement in Federal Funds rate would be due to employing the revised data. Sometimes, it would not be successful as much as described in Taylor (1993).

In sum, one reason for the criticism would be the rule could not provide enough micro-foundation about decision-making mechanism so that it would not be reasonable that a single policymaker makes the decision only based on output and inflation gap, ignoring the data availability and monetary policy committee decision-making mechanism. In other words, Taylor rule cannot explain which indicator they pay attention when there is economic crisis and how each indicator

---

2 For example, Ang and Piazzesi (2003) employed real activity and inflation factor data using factor analysis instead of inflation and GDP data, typically used in empirical analysis in Taylor rule.

3 Of course, the single agent problem, minimizing the welfare loss function, could be enough micro-foundation. But, it is too simple to capture the dynamic decision-making process in monetary policy committee.
can have forecasting error. In order to decrease the tension between Taylor rule and its criticism, this chapter employs the theory of team, which investigates how members in team interact to determine optimal decision. Using team theory, this chapter shows that how optimal monetary policy is resulted from committee members’ interaction, how policy disturbance can be seen as forecasting error, and how timeliness of data availability can be incorporated into decision-making process.

Figure 1.1: Deviations of actual funds rate from estimated value from Taylor rule (Rudebusch 2006)

1.2 Decision Making in the Committee

"My experience as a member of the FOMC left me with a strong feeling that the theoretical fiction that monetary policy is made by a single individual maximizing a well-defined preference function would be a wonderful simplification if it were true. It’s not true. ... [In the committee], there’s a variety of different views about how the economy works, how monetary policy should be conducted, and what the appropriate targets for monetary policy are. There’s a lot of give and take, a lot of discussion, and a lot of negotiation. And that’s how decisions are made."
misses something important. In my view, monetary theorists should start paying some attention to the nature of decision-making by committee, which is rarely mentioned in the academic literature.” (Blinder 1998, pp.22)

Even though size of the monetary policy committee, and decision making process (for example, simple majority voting, or consensus) are different to each central bank in each country, one important common factor in decision-making in central banks is the policy results from committee member interaction. The questions arise here are why committee decision-making is needed in contrast with the typical assumption, single agent decision-making, how committee makes the decision and what effect it has on. First, according to literatures, theoretically, monetary policy conducted by the committee is better than one by individual decision making in terms of the accuracy since the committee can have more information than single agent. (Gerlach-Kristen 2006) Meanwhile, empirically, group decision-making performs better than individual decision-making even though the speed of group decision-making is slower. (This is because the speed can decrease the accuracy due to the lag response.) (Blinder and Morgan 2006) In this vein, committee decision-making in monetary policy is the natural result.

Second one, most of the literature focuses on voting behavior and how voting rule affects on its final-decision. However, these do not focus on its general effect on the economy. In this point of view, needless to say, the study about committee decision-making is very important to study how they decide the optimal policy and how its mechanism affects on the economy. But, the complexity arises when there is the conflict of interest in committee decision making. In other words, if each member has the different objective in decision-making or perception in economic

---

5 Differences among the central banks is well described in Fujiki (2005).
6 In this vein, Ehrmann and Fratzscher (2006) showed that the committee in the central bank has the effect on financial market and this effect depends on how the committee organizes or who is most effective person through publishing booklet like economic outlook or press. However, this paper did not show direct effect of decision-making in the committee.
situation, they will do strategic voting behavior. Due to their strategic voting behavior, that is, they act individualistically, it is not easy to find out the closed-form of monetary policy rule. This is because the derived optimal policy is the outcome of game-theoretic results in this case, and this can be changed by the preference determined by which situation they face or which payoff they can receive when the decision is made.

According to Meade and Sheets (2005), Fed committee members focus on their regional situation more than national wide issues. This empirical result implies that there are some conflict of interest in FOMC. But, according to Yellen (2005), the former president and chief executive officer of Federal Reserve Bank of San Francisco, "In fact, I think FOMC members behave far less individualistically and strategically than assumed in some of the model....." (pp.2) In both points of view, it would be possible to conclude that the difference among members is not due to its conflict of interest but their different observation about economic situation or different perception. In other words, because they could perceive the situation differently due to regional difference, their voting behavior seem to be individualistic or strategic.

If, however, it is assumed that each member has the same objective, for example, price stabilization or minimizing output gap with different perceptions about economy, committee decision-making can be seen as a team decision-making. In this case, decision-making in the committee would be relatively simpler than game-theoretic situation in committee decision-making.

1.2.1 The theory of Team

In economics, typically individual agent decides own optimal decision to maximize or minimize their objectives. In line with this, the main property of the team is that there is single objective.
or payoff to all members so that they have the same objective or payoff. In terms of this, we can classify organizations based on following propositions and find out general proposition of team.\(^9\)
S will denote the set of all states.

a. For every \(i(i = 1, \ldots, n)\), there exists a complete ordering \(G_i\) on \(S\) [That is, every group member has a preference ordering]: Rationality of members.

b. There exists a transitive ordering \(G_0\) on \(S\) [That is, the states can be ordered for the group at least partially]: Transitivity of group interests.

c. For any \(s, s'\) if \(s G_i s'\) for all \(i = 1, \ldots, n\) then \(s G_0 s'\). [That is, if \(s\) is not worse than \(s'\) it is not worse than \(s'\) for the group itself.]: The so-called Pareto Optimality principle.

d. For any \(s, s'\) in \(S\), \(s G_0 s'\) or \(s' G_0 s\). [That is, for no pair of states is it impossible to say which is preferable for the group, unless the two are equally desirable]: Completeness of group preferences. (Note that B and D imply together complete ordering.)

e. For all \(i = 1, \ldots, n\), \(s G_i s'\) if and only if \(s G_0 s'\). [That is, all individuals have identical interest among themselves and with the group]: Solidarity.

From above propositions, (1) if a is satisfied we call n group members rational, (2) if a, b, and c are satisfied, the group is coalition, (3) if a, b, c and d are satisfied we call it a foundation, (4) if a, b, c, d and e are satisfied we call it a team.\(^{10,11}\) Therefore team can be defined as organization in which each member controls different action and different information but have same objectives, that is, there is a single goal or payoff to all members in the team.

\(^9\)For details about the team theory, see Marschak (1955) and McGuire and Radner (1972). For its application, see Beckmann (1958) and McGuire (1961).\(^{10}\) These propositions are quoted from Marschak (1960).

\(^{11}\)In case of coalition or foundation, its final-decision depends on game-theoretic outcome. Therefore, in order to understand its outcome, we need to investigate which solution concept is applied to get final-decision, for example, Hurwicz solution (Hurwicz 1951) or Minimax solution. As mentioned above, because the payoff of each member in game situation depends on their own situation, in other words, because it is changed as time passed by due to the difference in their objective, this leads individualistic behavior and general solution for decision making in organization is not easy.
Typically, team-theoretic decision problem has 5 ingredients, random variable, observation, decision variable, strategy (decision rule), and loss criteria.\footnote{These are quoted from Ho (1980)}

a. States of Nature: $\xi = [\xi_1, \cdots, \xi_m]$. They may be measurement noise, random disturbance, uncertain initial conditions, etc.

b. A set of Observations: $z = [z_1, \cdots, z_n]$ which are given functions of $\xi$, i.e., $z_i = \eta_i(\xi_1, \cdots, \xi_m), i = 1, \cdots, n$. In general, $z_i$ is known as the information or observation available to the $i$th decision maker. The set $\{\eta_i | i = 1, \cdots, n\}$ is denoted as the information structure of the problem.

c. A set of decision variables: $u \equiv [u_1, \cdots, u_n]$

d. The strategy (decision rule, control law) of the $i$th decision maker is a map $\gamma_i : Z_i \rightarrow U_i, u_i = \gamma_i(z_i)$

e. The loss (payoff) criterion of the problem is a map $L : \Xi \times U \rightarrow R$, i.e., $Loss = L(u_1, \cdots, u_n, \xi_1, \cdots, \xi_m)$

From these, the team decision can be defined as follows; Find $\gamma_i \in \Gamma_i, \forall i$ in order to minimize $J = E_{\xi}[L(u = \gamma(\eta(\xi)), \xi)]$ or $Min_{\gamma \in \Gamma} J(\gamma)$. Generally, in team theory, typical payoff function is assumed as a quadratic cost function. This is because its optimal decision for a quadratic function is sometimes linear and easy to handle.\footnote{See Radner (1962)} Due to this reason, common goal for team in this also uses a quadratic cost function.

$$J(\gamma_1, \cdots, \gamma_N) = E[3] = E[\frac{1}{2} u^T Q u + u^T S \xi + u^T c]$$

\footnote{These are quoted from Ho (1980)}
Where $u = [u_1, \cdots, u_N]'$, $Q$ is symmetric positive definite, $u_i$ are given by above, and the expectation is taken with respect to a priori $\xi$. Given the quadratic cost function, optimal decision is not partial derivatives with respect to control variable, $u_i$. As mentioned above, because optimal decision for team is the result of interaction, optimal decision has to incorporate individual optimal decision into its final decision $f$. Therefore, to find out optimal team decision, we need the additional criteria as follows\(^{14}\):

\[
J(\gamma_1^*, \cdots, \gamma_{i-1}^*, \gamma_i, \gamma_{i+1}^*, \cdots, \gamma_N^*) \leq J(\gamma_1^*, \cdots, \gamma_{i-1}^*, \gamma_i^*, \gamma_{i+1}^*, \cdots, \gamma_N^*)
\]  

For all $\gamma_i \in \Gamma$ and for all $i$

Simply this inequality implies optimal team decision has to be optimal to all members in the team. But the problem is that each individual optimal could not be team optimal. In other words, because this inequality includes non-cooperative individual decision in team, this condition could not be team optimality. In this vein, this is necessary for global optimal but not sufficient condition. Therefore we call this *member-by-member optimal*.

For this reason, we rewrite (1.1) as

\[
J = E\{3[\gamma_1^*(z_1), \cdots, \gamma_{i-1}^*(z_{i-1}), \gamma_i(z_i), \gamma_{i+1}^*(z_{i+1}), \cdots, \gamma_N^*(Z_N), \xi]\}
\]

\[
= E_{z_i}\{E[3|z_i]\} \text{ for fixed } \gamma_1^*, \cdots, \gamma_{i-1}^*, \gamma_i^*, \gamma_{i+1}^*, \cdots, \gamma_N^*\]

Where second is conditional expectation given any values of $z_i$. This implies that optimal team decision is one in which team optimal decision is made based on other members’ previous decisions so that the team problem is to find out the optimal decision given fixed other members’ deci-

\(^{14}\)The necessary condition for optimality is well described in [Ho and Chu (1972)]
sions (control law), $\gamma_1^*, \cdots, \gamma_{i-1}^*, \gamma_{i+1}^*, \cdots, \gamma_N^*$ Therefore we can rewrite common goal for team as follows:

$$
\min_{u_i} E\left[ \frac{1}{2} u^T Q u + u^T S \xi + u^T c | z_i \right] \equiv \min_{u_i} J_i
$$

where $u_i^T = [\gamma_i^T(z_1), \cdots, \gamma_i^T(z_{i-1}), \gamma_i^T(z_i), \gamma_i^T(z_{i+1}), \cdots, \gamma_i^T(z_N)]$ for fixed $\gamma_j^T(z_j), j \neq i$, and any $z_i$. In this case, a necessary condition for the optimality of team could be obtained by partial derivatives of $J_i$ with respect to $u_i$. But, this derivative must be considered with communication pattern. This is because previous decisions $u_j$ are included in $u_i$ through communication. Finally, communication patterns are very important to get the optimality of team.

### 1.2.2 Communication Pattern in a Team

One of most important factor in decision-making is information. Simply, because information would describe the situation about economy, it could be coming out from data-generating process. After observing economic situation through economic indicators, for example, inflation index, GDP, or quantity of money, decision-maker can get some information describing the situation. In contrast with this, one property of information is sharable. That is, by communication, decision makers can update their information which is based on others' observations even with some observational

---

15 The main reason why it assumes that other members’ previous decisions are made before final decision is typically decision making in team is not made simultaneously except decision-making in voting system. In other words, in team, because enough discussion is made before final decision-making, each member would know others’ decisions. Therefore it is natural to interpret the team problem as above. However, communication pattern among members is additional problem in team decision problem.

16 Because of interactions via communication, the partial derivative with respect to $u_i$ is $Q_i \gamma_i + \sum_{j \neq i} Q_{ij} E(\gamma_i | z_i) + S_i E(\xi | z_i) + c_i + \sum_{j \neq i} [u_i^T Q_{ij} \frac{\partial}{\partial u_i} E(\gamma_j | z_i) + \frac{\partial}{\partial u_i} E(\gamma_j^T S_j \xi | z_i) + \frac{\partial}{\partial u_i} E(\gamma_j^T c_j | z_i) + \sum_{k \neq i} \frac{\partial}{\partial u_i} E(\gamma_k^T Q_{kj} \gamma_j | z_i)] = 0$. For details, see [Ho and Chu (1972)]

17 In here, the problem how much economic agents pay attentions to get information is arise, that is, rational inattention problem proposed by [Sims (2003)] in this paper, rational inattention problem is ignored because it is not reasonable assumption that central bank do not pay full attention to observe economic situations.

18 For example, FOMC is reported Greenbook for monetary policy.
error. Therefore, under the team situation in which there are many decision makers, communication pattern is very important matter to determine how they update the information about economic situation, and how they make the decision based on updated information through communication. Bavelas (1951) showed that efficiency of decision-making is different among communication patterns, that is, by experimenting some communication patterns, "same task-forced but different communication pattern group" has different efficiency due to different updating in information through communication. Therefore communication patterns in team is very important to the decision-making.

If it assumes that there is final-decision maker in team and all communication is finally heading to final-decision maker, for example, B in figure 1.2, its efficiency of team is depending on how they communicate, that is, how they are organized to communicate. Needless to say, based on above assumptions, if there is only one final-decision maker in team and final-decision maker is placed on left-end, communication pattern will be headed from right-end to left-end, vice versa. In both case, since individual observation on left-end individual or right-end individual is different to each other, intuitively its decision has different efficiency. Hence, it concludes that information structure depends on its communication pattern and communication is important one to determine its efficiency in decision-making since information is differently updated by communication.

In sum, information structure can be defined as follows, if communication exists in team:

$$z_i = H_i \xi + \sum_j D_{ij} u_j, \forall i.$$  

19 In communication theory, put simply, communication can be measurable by how much information can be contained in communication. In this point of view, Shannon and Weaver (1963) suggested "Entropy" to measure the information in communication. In this paper, it is assume that all information are transferable to the other agent within team to ignore entropy.

20 Of course, as in voting system, all constituent individual of team can be final-decision maker. In this case, instead of final-decision maker, decision maker on final-stage is proper one to describe the situation.
where \( H_i \) and \( D_{ij} \) are some positive matrices. Typically, it can be interpreted as weighting on their information and decisions. In above equation, as mentioned above, information structure consists of one’s own observation and others’ decisions from their own observation. In other words, second component on RHS is updated information through the communication among team members.

In dynamic team\(^{21}\) situation as described above, for example, information structure in D in figure 1.2 is as follows\(^{22}\):

\[
\begin{align*}
z_1 &= H_1 \xi \\
z_2 &= H_2 \xi \\
z_3 &= H_3 \xi \\
z_4 &= H_1 \xi + D'_{41}u_1 + D'_{42}u_2 + D'_{43}u_3 \\
z_5 &= H_1 \xi + D'_{51}u_1 + D'_{52}u_2 + D'_{53}u_3 + D'_{54}u_4
\end{align*}
\]

From above, it can conclude that \( z_5 \) or \( z_4 \)'s information structure can deduce from communicating the decision of the precedent. But, the problem to solve optimal decision given information structure is that the dependence among members due to the communication make the solution non-linear. In

\(^{21}\)In team theory, there are two types of team problem. One is a static-team problem, the other one is a dynamic-team problem. The difference between problems relies on whether there exists communication or not. Simply, in static-team problem, typical information structure is \( z_i = H_i \xi, \forall i \).

\(^{22}\)For more examples, see Ho and Chu (1972).
other words, because the information structure depends on decisions of his or her precedents, that is, decisions made in lower rank affects on the dependence through information structure, optimal decision of some members could not be linear. For instance, if one member has non-linear decision-making, then technically, final-decision also would not be linear one. In this case, it is not easy to find out unique final optimal decision of team due to its dependence. In this difficulty, Chu (1972) suggested an alternative method to find out the optimal solution. This is summarized in figure 1.3.

Figure 1.2: An example of communication patterns (Bavelas 1951)

1.2.3 The Optimal Monetary Policy under Team Situations

In the United State, monetary policy is determined by, as known, Federal Reserve System, especially, the Borad of Governors of the Federal Reserve System and Federal Open Market Committee (hereafter FOMC). The Board of Governors of the Federal Reserve System is responsible for the discount rate and reserve requirement while FOMC is responsible for open market operations. Even Witsenhausen (1968) showed that generally optimal decision of members may not be linear in information.
with different policy instruments between them, FOMC is the most important policymaking body of the Federal Reserve System. This is because the FOMC consists of the five member of the Board of Governor\textsuperscript{24} and five Reserve Bank Presidents. Therefore, the decision made at the FOMC would reflect the decision in the Board of Governor\textsuperscript{25}

In FOMC, decision-making process is as follows:

1. System staff prepares reports on past and prospective.

2. Staff officer gives oral presentation on current and prospective situation in the meeting.

3. After these report, reports and oral presentation, each member in FOMC expresses his or her own views on the state of the economy, forecast for the future, and appropriate policy.

\textsuperscript{24}The Board of governors are composed of 7 members appointed by the President and confirmed by the Senate

\textsuperscript{25}For details, see Boyer (2005)
4. And then, each recommends more explicit policy for the coming inter-meeting period.

5. Finally, the committee has to reach a consensus about appropriate policy or course for policy.

In sum, FOMC decision-making could be described by that after each expresses their opinions about the economy and states own policy recommendation, final decision is made. In order to resemble the decision-making process in FOMC closely, it is assumed here that there are three members in the committee under team situation. Two are members in FOMC and one, final decision-maker, is the president in FOMC. Since final decision is made after stating or expressing policy recommendation, communication pattern in this situation would be described as in figure 1.4.

This is because other members’ opinions are some kind of others’ decision-making based on their information. Put simply, this artificial FOMC could be seen as one in which two agents update or report their information (or decisions based on observation) to their headquarter.

Figure 1.4: The communication pattern in committee decision-making under artificial FOMC.

Therefore, the information structure in this decision-making process is

\[ z_1 = H_1 \xi \quad \quad z_2 = H_2 \xi \]
\[ z_3 = \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} \xi + \begin{bmatrix} 0 \\ 0 \\ D'_{31} \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ 0 \\ D'_{32} \end{bmatrix} u_2 \]

### 1.2.4 Solution Technique

#### General framework

One of results in a team problem is that a dynamic-team problem can be solved by the solution of a static-team problem obtained by Radner (1962). This result can simplify the problem in a dynamic-team even with some limitation.

Suppose that a team problem is the problem in static-team. In this case, each member’s information structure is a linear function of \( \xi \), where \( \xi \) is Gaussian probability density function, \( N(0, X) \).

\[ z_i = H_i \xi_i \]

where \( H_i \) is \( q_i \times n \) with \( n > q_i \). This is because there is no communication among members. Given the common goal for the team,

\[
E\left[ \frac{1}{2} u^T Q u + u^T S \xi + u^T c \bigg| z_i \right]
\]

The optimal solution for this is

\[
Q_{ii} \gamma_i + \sum_{j \neq i} Q_{ij} E(\gamma_i \mid z_i) + S_i E(\xi \mid z_i) + c_i = 0
\]

---

\[26\] This subsection and following subsections are re-summarized from [Ho and Chu (1972) and Chu (1972)].
Compared to the dynamic-team, this is more simplified one. If the control law is $u_i = \gamma_i(z_i) = A_i z_i + b_i$, $\forall i$ (where $A_i$ and $b_i$ are $k_i \times q_i$ matrix and $k_i$ vector, respectively), then optimal solution is as follows:

$$Q_{ii}(A_i z_i + b_i) + \sum_{j \neq 1} Q_{ij} E(A_j H_j \xi + b_j | z_i) + S_i E(\xi | z_i) + c_i = 0$$

Because $z_j = H_j \xi$

$$Q_{ii}(A_i z_i + b_i) + \left[ \sum_{j \neq 1} Q_{ij} E(A_j H_j + b_j | z_i) + S_i \right] E(\xi | z_i) + c_i = 0$$

Since $E(\xi | z_i) = \frac{\text{Cov}(\xi, z_i)}{\text{Var}(z_i)} = X H_i^T (H_i X H_i^T)^{-1} \cdot z_i$, where $X$ is the variance of $\xi$, the above equation is rewritten as

$$\left( \sum_{j=1}^N Q_{ij} b_j + c_i \right) + \left[ Q_{ii} A_i + \left( \sum_{j \neq 1} Q_{ij} A_j H_j + S_i \right) \cdot X H_i^T (H_i X H_i^T)^{-1} \right] z_i = 0, \forall z_i \text{ and } \forall i$$

Since the control law is $u_i = \gamma_i(z_i) = A_i z_i + b_i$, $(b_1^T, b_2^T, \cdots, b_N^T) = c^T Q^{-1}$ from first component in above equation is obtained, and

$$Q_{ii} A_i + \sum_{j \neq 1} Q_{ij} A_j H_j \cdot X H_i^T (H_i X H_i^T)^{-1} = -S_i \cdot X H_i^T (H_i X H_i^T)^{-1}$$

or

$$\sum_{j \neq 1} Q_{ij} A_j H_j X H_i^T = -S_i X H_i^T$$

Therefore, it can conclude that $u_i = A_i Z_i + b_i$ is optimal for the static-team problem. From this result, Radner (1962)'s theorem is obtained, that is, in static-team problem, the control law,
\( u_i = A_i z_i + b_i \), is optimal. This implies that the static-team problem with a quadratic objective function has linear law of control.

Even though the solution in static-team problem is tractable one, this cannot be applied to the problem in a dynamic-team. As mentioned above, the optimal solution in dynamic-team may not be linear due to the communication. Due to this reason, Ho and Chu (1972) proposed the special information structure, "Partially nested information structure." The partially nested information structure is one in which each decision-maker in a team can always deduce the action of information of its precedents since all decisions and information are updated to the followers. Under partially nested information structure, the control law and information of all precedents is fixed. For example, as shown above, the information structure for the final decision maker in artificial FOMC is

\[
    z_3 = \begin{bmatrix}
        H_1 \\
        H_2 \\
        H_3 
    \end{bmatrix} \xi + \begin{bmatrix}
        0 \\
        0 \\
        D'_{31}
    \end{bmatrix} u_1 + \begin{bmatrix}
        0 \\
        0 \\
        D'_{32}
    \end{bmatrix} u_2
\]

Since all decisions and information is updated to the final decision maker, \( \sum D_{ij} u_j \) is redundant one. In other words, because the decision of precedents is deduced from the information given to the previous decision-maker, \( \sum D_{ij} u_j \) can be deleted from \( z_3 \). (where \( j \) is a precedent.) Let \( \tilde{z}_3 \) be the partially nested information structure. Then, \( z_3 \) can be modified as follows:

\[
    \tilde{z}_3 = \begin{bmatrix}
        H_1 \\
        H_2 \\
        H_3 
    \end{bmatrix} \xi
\]

As seen above, this is exactly same information structure in static-team. Based on this, following
theorem is obtained:

*Theorem 1* (Ho and Chu [1972], pp.20) In a dynamic team with partially nested information structure, 
\[ z_i = H_i \xi + \sum_j D_{ij} u_j, \]
is equivalent to an information structure in static form for any fixed set of control laws, \( \tilde{z}_i = \{H_i \xi | j \{i \text{ or } j = i\}\}. \) This theorem implies that \( z_i = \tilde{z}_i = H_i \xi. \) Finally, it can be conclude that in a dynamic-team with partially nested information structure, the optimal control for each member exist, is unique and linear in \( z_i. \)

Under the dynamic-team problem with partially nested information structure is simplified into the static-team problem. However, since partially nested information structure is one special case in a team problem, it would not be general to apply this information structure to general dynamic-team problem. This is because the information structure in a dynamic-team is not always partially nested. As mentioned above, moreover, since Witsenhausen (1968) argued that generally optimal action of each member in a team might not be linear in information, more general approach is needed to solve the dynamic-team problem.

**Equivalent controls and construction of a team problem with partially nested information (the auxiliary problem)**

Due to the problem mentioned above, Chu (1972) suggested the alternative way to solve a dynamic-team problem with non-nested information structure. To solve general dynamic-team problem, Chu (1972) invented two concepts, equivalent control and construction of the auxiliary problem. Intuitively, this is the indirect solution concept for a dynamic-team.

Suppose that a dynamic-team problem with non-nested information structure and one with partially nested information structure have the optimal law of control. (Let \( \Omega \) and \( \tilde{\Omega} \) denote a dynamic-team problem with non-nested information structure and one with partially nested information structure, respectively.) Regardless of which team problem is chosen, the optimal law of controls is \( u_i = \)

\(^{27} \textit{Theorem 2} \) (Ho and Chu [1972], pp.20)
Let \( u_i \) and \( \tilde{u}_i \) denote the optimal law of control in \( \Omega \) and \( \tilde{\Omega} \), respectively. The information structure in \( \Omega \) is \( z_i = \eta_i(\xi, u_{i-1}) \). Likewise, the information structure in \( \tilde{\Omega} \) is \( z_i = \eta_i(\xi, \tilde{u}_{i-1}) \).

Since the information function, \( z_i \) (or \( \tilde{z}_i \)) is mapping from the information set, \( \Xi \), and the set of the law of control, \( U_{i-1} \) (or \( \tilde{U}_{i-1} \)) to the information set, \( Z_i \) (or \( \tilde{Z}_i \)) and the control function, \( \gamma_i \) (or \( \tilde{\gamma}_i \)), is mapping from \( Z_i \) (or \( \tilde{Z}_i \)) to \( U_i \) (or \( \tilde{U}_i \)), \( u_i \) is the composite function for \( \gamma_i \eta_i \) such that \( u_i = \gamma_i[\eta_i(\xi, u_{i-1})] = \gamma_i\eta_i(\xi, u_{i-1}) \) (Likewise, \( \tilde{u}_i = \tilde{\gamma}_i[\tilde{\eta}_i(\xi, \tilde{u}_{i-1})] = \tilde{\gamma}_i\tilde{\eta}_i(\xi, \tilde{u}_{i-1}) \)). Therefore, the composite functions, \( p_i(\xi) \) (or \( \tilde{p}_i(\xi) \)) can be defined as one mapping from \( \Xi \) to \( U_i \) (or \( \tilde{U}_i \)). This can be described in figure 1.5.

![Figure 1.5: Equivalent Controls](Chu1972)

After constructing the composite function, \( p_i \) and \( \tilde{p}_i \), it can be conclude that \( p_i = \tilde{p}_i \) if and only if \((\eta_1, \cdots, \eta_N; \gamma_1, \cdots, \gamma_N)\) and \((\tilde{\eta}_1, \cdots, \tilde{\eta}_N; \tilde{\gamma}_1, \cdots, \tilde{\gamma}_N)\), the control law \((\gamma_1, \cdots, \gamma_N)\) and \((\tilde{\gamma}_1, \cdots, \tilde{\gamma}_N)\) are equivalent. Therefore, if \( p_i(\xi) = \tilde{p}_i(\xi) \), then \( J = E[J(\xi, U_N)] = E[J(\xi, p_N(\xi))] = E[J(\xi, \tilde{p}_N(\xi))] = E[J(\xi, \tilde{U}_N)] = \tilde{J} \) (In here, \( J \) is the objective function in a team problem, therefore, \( E[J(\cdot)] \) is the payoff in a team.) This implies that both \( \Omega \) and \( \tilde{\Omega} \) have same payoff if and only if \( p_i(\xi) = \tilde{p}_i(\xi) \).

Suppose that both problems have the same payoff. However, the payoff of \( \Omega \) is bound below by \( \tilde{\Omega} \) since the partially nested information structure has more information than non-nested one.

\(^{28}\) Theorem 1 (Chu1972, pp.23)
In order to find $\gamma_1, \cdots, \gamma_N$ for $\Omega$ such that resulting $p_i(\xi)$ is equal to $\tilde{p}_i(\xi)$, Chu (1972) proved following theorem and corollary:

**Theorem 2** (Chu 1972, pp. 24): Define $\tilde{p}^*_i$ as the composite control corresponding to the optimal control $\tilde{\gamma}^*_i$ of $\tilde{\Omega}$ for each $i$; and define functions

$$g_i(\xi) = \eta_i(\xi, \tilde{p}_1^*(\xi), \tilde{p}_2^*(\xi), \cdots, \tilde{p}_{i-1}^*(\xi)), \forall i$$

If there exists some functions $r_i$ from $Z_i$ to $U_i$ such that $\tilde{p}^*_i = r_i g_i, \forall i$, then $(\gamma_1, \cdots, \gamma_N)$ is optimal for problem $\Omega$ and they are equivalent to $(\tilde{\gamma}^*_1, \cdots, \tilde{\gamma}^*_N)$ of the problem $\tilde{\Omega}$.

Suppose that $g_i$ is invertible for all $i$. By theorem 2, $\gamma_i = \tilde{p}_i^* g_i^{-1}$ for all $i$. Therefore, $p_i(\xi) = \gamma_i \eta_i(\xi) = r_i \eta_i(\xi) = \tilde{p}_i^*$. By the recursive, it can be obtained that $p_N = \gamma_N \eta_N(\xi; p_1(\xi), \cdots, p_{N-1}(\xi)) = r_N \eta_N(\xi, \tilde{p}_1(\xi), \cdots, \tilde{p}_N(\xi)) = [\tilde{p}_N^* g_N^{-1}] g_N(\xi) = \tilde{p}_N^*(\xi)$. Following corollary can be obtained:

**Corollary** (Chu 1972, pp. 24): If $g_i$ is invertible for all $i$, the optimal control laws for problem $\Omega$ can be found and they are equivalent to those of problem, $\tilde{\Omega}$.

In sum, if the dynamic-problem with the linear law of control, then the solution of this problem can be obtained through the solution of the static-team problem by Radner (1962)'s theorem. However, partially nested information structure is not general information structure in dynamic-team. Due to this problem, Chu (1972) constructed the auxiliary problem, that is, the dynamic-team problem with partially nested information structure. If both the dynamic-team problem with partially nested one and one with non-nested one have the same payoff using the composite control function, $p_i$ and $\tilde{p}_i$, by theorem 1 in Chu (1972), both have the equivalent control. In this case, the dynamic-team with non-nested information structure can be solved by the solution of the static-team. As shown above, however, some special function ($r_i$) is needed for $p_i$ to be equal to $\tilde{p}_i$. Moreover, this
procedure cannot guarantee that both team problems have the same payoff even if the composite control function in each problem is obtained. (See figure [1.3]

1.2.5 Numerical Solution to Derive the Optimal Monetary Policy

Under this information structure, the optimal monetary policy can be derived as follows:

Step 1. Construct auxiliary (partially nested) information structure.

\[ \tilde{z}_1 = z_1 = \xi \quad \tilde{z}_2 = z_2 = \xi \]

\[ \tilde{z}_3 = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} \xi \\ \xi \\ u_1 + u_2 + \xi \end{bmatrix} \]

Because information is updated to final decision maker, that is, partially nested information structure is assumed, redundant information can be deleted. In this case, information structure in a dynamic-team is same one in a static-team. Therefore, equivalently, optimal decision in a dynamic-team is

\[ \hat{z}_1 = \xi \quad \hat{z}_2 = \xi \quad \hat{z}_3 = \xi \] (1.6)

For simplicity, it is assume that

\[ Q = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad S = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \]

\[ C = 0 \quad Pr(\xi) = N(0, \sigma^2) \]

For more examples, see Chu (1972)
The optimal solution of auxiliary problem given auxiliary information structure is

\[
\begin{bmatrix}
\tilde{u}_1 \\
\tilde{u}_2^* \\
\tilde{u}_3^*
\end{bmatrix} = -Q^{-1} S \xi =
\begin{bmatrix}
3/4 & -1/4 & -1/4 \\
-1/4 & 3/4 & -1/4 \\
-1/4 & -1/4 & 3/4
\end{bmatrix}
\begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix} \xi = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} \xi
\]

or

\[
\tilde{u}_1 = -\frac{1}{4}\tilde{z}_1, \quad \tilde{u}_2^* = -\frac{1}{4}\tilde{z}_2, \quad \tilde{u}_3^* = -\frac{1}{4}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\tilde{z}_3 \tag{1.7}
\]

From (7), composite control corresponding to the optimal control for auxiliary problem is

\[
\tilde{p}_1^*(\xi) = \tilde{p}_2^*(\xi) = \tilde{p}_3^*(\xi) = -\frac{1}{4}\xi \tag{1.8}
\]

and

\[
\tilde{J}^* = E \begin{bmatrix}
\frac{1}{2} \begin{bmatrix}
\tilde{u}_1^* & \tilde{u}_2^* & \tilde{u}_3^*
\end{bmatrix}
& \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}
& \begin{bmatrix}\tilde{u}_1^* \\ \tilde{u}_2^* \\ \tilde{u}_3^* \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
& \begin{bmatrix} \tilde{u}_1^* \\ \tilde{u}_2^* \\ \tilde{u}_3^* \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
& \begin{bmatrix} \tilde{u}_1^* \\ \tilde{u}_2^* \\ \tilde{u}_3^* \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
\end{bmatrix}
\]

\[
= E \begin{bmatrix}
\frac{1}{2} \begin{bmatrix} -\frac{1}{4}\tilde{z}_1 & -\frac{1}{4}\tilde{z}_2 & -\frac{1}{4}\tilde{z}_3 \end{bmatrix}
& \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}
& \begin{bmatrix} -\frac{1}{4}\tilde{z}_1 \\ -\frac{1}{4}\tilde{z}_2 \\ -\frac{1}{4}\tilde{z}_3 \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
& \begin{bmatrix} -\frac{1}{4}\tilde{z}_1 \\ -\frac{1}{4}\tilde{z}_2 \\ -\frac{1}{4}\tilde{z}_3 \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
& \begin{bmatrix} -\frac{1}{4}\tilde{z}_1 \\ -\frac{1}{4}\tilde{z}_2 \\ -\frac{1}{4}\tilde{z}_3 \end{bmatrix}
& \begin{bmatrix} 1 \end{bmatrix}
\end{bmatrix}
\]
\[
E \left[ \frac{1}{2} \xi^2 \begin{bmatrix} -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{bmatrix} + \xi^2 \begin{bmatrix} -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right] = E \left[ \frac{1}{2} \xi^2 \begin{bmatrix} -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} - \frac{3}{4} \xi^2 \right] = E \left[ \frac{1}{2} \xi^2 \frac{3}{4} - \frac{3}{4} \xi^2 \right] = \frac{3}{8} \sigma^2 - \frac{6}{8} \sigma^2 = \frac{3}{8} \sigma^2
\]

Step 2. Substitute \( p_i^*(\xi) \) into \( z_i \)

\( g_1(\xi) = z_1 = \xi \)

\( g_2(\xi) = z_2 = \xi \)

\( g_3(\xi) = \xi + \tilde{p}_1^*(\xi) + \tilde{p}_2^*(\xi) = \frac{1}{2} \xi \)

By corollary, the optimal controls for team are

\[
\begin{align*}
    u_1^* &= \gamma^*(z_1) = \tilde{p}_1^* g_1^{-1}(z_1) = -\frac{1}{4} z_1 \\
    u_2^* &= \gamma^*(z_2) = \tilde{p}_2^* g_2^{-1}(z_2) = -\frac{1}{4} z_2 \\
    u_3^* &= \gamma^*(z_3) = \tilde{p}_3^* g_3^{-1}(z_3) = -\frac{1}{4} \left( \frac{2}{3} \right) z_3 = -\frac{1}{2} z_3
\end{align*}
\]
and the optimal payoff function $J^*$ corresponding to above is

$$J^* = E \left[ \frac{1}{2} \xi^2 \left[ -\frac{1}{4} - \frac{1}{4} - \frac{1}{2} \right] \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} + \frac{1}{4} \xi \left[ -\frac{1}{4} - \frac{1}{4} - \frac{1}{2} \right] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right]$$

$$= E \left[ \frac{1}{2} \xi^2 \left[ -\frac{5}{4} - \frac{5}{4} - \frac{5}{4} \right] - \frac{1}{4} \xi \left[ -\frac{1}{4} - \frac{1}{4} - \frac{1}{4} \right] \right]$$

$$= E \left[ \frac{1}{2} \xi^2 \frac{5}{4} - \xi^2 \right] = \frac{5}{8} \sigma^2 - \frac{8}{8} \sigma^2 = -\frac{3}{8} \sigma^2$$

Because both $\tilde{J}^*$ and $J^*$ have same value given control variables, the optimal decision in this team situation is

$$u_3^* = \frac{1}{2} z_3 = \frac{1}{2} \left[ u_1^* + u_2^* + \xi \right] = \frac{1}{2} \left[ -\frac{1}{4} \xi + \frac{1}{4} \xi + \xi \right] = \frac{1}{4} \xi$$

(1.9)

As mentioned above, since $\xi$ is known as the state of nature to the decision maker, simply optimal decision is function of observations based on the state of nature.\(^{30}\) In other words, optimal decision is

$$u_3^* = \frac{1}{4} \eta(\xi) = \frac{1}{4} z.$$  This implies that optimal monetary policy depends on how committee members estimate the state of nature with imperfect observation of the event or how they extract the signal of the event. Therefore optimal decision is the making decision under imperfect observation or signal extraction problem.\(^{31}\) For example, suppose that the observation, $z$, is realization of a linear

\(^{30}\)In control theory, under some assumptions, the optimal control problem could be solved by the optimal observation problem, or optimal estimation problem. Typically, this is known as the separation principle. If we assume that there is asymmetric information problem between central bank and private agents, for example, the situation in Barro and Gordon (1983), the separation principle does not hold. Especially, according to Aoki (2003), the optimal estimation or optimal observation is depending on choice of policy if there is the failure of the separation principle.

\(^{31}\)According to Swanson (2004), certainty-equivalence problem is arise in case of signal extraction problem. In this
function of $\xi$ or law of motion of state of nature. Thus,

$$z = A\xi$$

where $A$ is a known coefficient matrix of appropriate dimension to the state of nature. Typically, as in Taylor rule, observation could be combination of two elements in the state of nature,

$$z = a_1\xi_1 + a_2\xi_2$$  \hspace{1cm} (1.10)

In (10), underlying state of the economy, $\xi$, would be policymakers’ optimal estimation, $\hat{\xi}$. This implies that

$$z = a_1\hat{\xi}_1 + a_2\hat{\xi}_2 + \nu$$  \hspace{1cm} (1.11)

where $\nu$ is the persistent revision or noise or forecasting error. (In here, policy disturbance can be seen as forecasting error using real-time estimate.) Therefore, if it is assumed that the policymakers use output gap and inflation gap as indicator variable for the economy

$$z = a_1(y - y^*) + a_2(\pi - \pi^*) + \nu$$  \hspace{1cm} (1.12)

Finally, optimal monetary policy in team situation is, (In here, the decision variable is the interest case, as uncertainty surrounding the economy or uncertainty about indicator variable is increased, coefficients of policy rule responded to the economy is attenuated.
rate.\footnote{Typically, $y^*$ is assumed as natural output. But, empirically, it can not be measurable. Due to this problem, in some case, instead of $y^*$, $\hat{y}^*$, estimation of natural output, is used. As mentioned in Swanson (2004), the former is signal extraction problem, the latter is imperfect observation problem. In order to avoid the problem of certainty-equivalence, it is assumed that $y^* = \hat{y}^*$}

\[
  r = \frac{1}{4}[a_1(y - y^*) + a_2(\pi - \pi^*) + \nu]
\]  

(1.13)

As seen before, this optimal monetary policy is consistent with \textit{Taylor} (1993). The only difference between these two is this rule is based on team decision-making process, whereas Taylor’s is not. In addition, optimal rule under team situation is not restrictive to add additional or different variables on Taylor compared to single agent problem. This is because team situation depends on how they perceive the situation through the economic indicators so that the observation could be combination of different kind of elements. Therefore, many different type of policy rule could be derived without amending the Fed’s objective function.

1.3 Conclusion - The optimal monetary policy under team situation

According to \textit{Goodfriend} (1991), Federal funds rate has been adjusted when there is new information updated to Fed.\footnote{The property of interest-rate targeting is well described in \textit{Goodfriend} (1991)} In this vein, the optimal policy decision-making could be a signal extraction problem since the information signals what happens in the economy. In this chapter, using team theory, optimal policy decision-making is simplified by signal extraction problem under the condition in which the separation principle holds. Even with very different micro-foundation, it is found that Taylor rule (\textit{Taylor} 1993) is still satisfied under the different optimization mechanism, committee decision-making.

In addition to this, this chapter also provides that optimal monetary policy depends on how individ-
ual decision-maker in the committee observes the economic situation through economic indicators. This implies that various type of Taylor rule is available based on the observations. Put simply, optimal monetary policy can be modified by which information they will choose or which indicators they will use to update their information sets. According to Rudebusch (2006), the deviation of Taylor rule from actual funds rates could be explained by macroeconomic developments, such as credit crunch, inflation or deflation scare. This implies that Federal Reserve appears to respond to not only current output gap and inflation gap but also additional information or indicators about the development.\footnote{Many proxies for financial market stability is used to explain.} In other words, the Taylor rule with only 2 components would be an inaccurate description of the actual monetary policy.

Even with this success, one weakness of the analysis is that, due to its complexity of team decision-making or committee decision-making, each member in the committee is assumed to be a rational observer. In other words, each agents’ observing the event has same variance. This implies that all agents in a team transmit the information without strategic action. If some agents transmit the information strategically, they will communicate with or report noisy information even though they observe or have the same information. Therefore, assuming the same variance cannot be appropriate.\footnote{Related to this, herding behavior in forecasting, is one to be able to explain why each agent has similar observation or forecasting under uncertain situation. According to Laster, Bennett and Geoum (1999) "if all forecasters have similar information and seek accuracy-consistent with the demands of intensive [forecasting data] users-their projection will cluster tightly around a consensus." (pp.298) In constrats with this, there is the deviation from consensus if forecasting agent has any incentive to deviate from that, for example, reputation or wages. The former effect is called "far-to-fall" effect, the latter is called "tighter priors" one. Because the magnitude of each effect depends on which industry forecasting agents are belong to, it is easy to determine whether forecasting or expectation is herding or not. For detail, see Lamont (2002).}
Chapter 2

Fool in the shower

2.1 Introduction

"The first time to turn on the water in the shower, needless to say, it is cold. Foolishly, the fool turns the knob to hot, and he gets too hot. So he turns knob to cold. He is doing this again and again, finally he only gets scalded when the water heats up with predicted lag"
time for each member to recognize the current economic situation as a boom or sluggish market because it takes time to gather the data to check what is going on the economy. Administrative lags suggest that it takes time to make or announce the decision for appropriate stabilization policy. In contrast with these, response lags is due to how much time it takes for stabilization policy to be effective in the economy. Among three lags, the main difference is that recognition lags and response lags are internal lags while response lags are external lags. Therefore, based on internal lags, it is reasonable to assume that there are internal lags when the committee decides its optimal decision-making.

In line with this, many researchers, especially monetary policy economists, try to incorporate this kind of lags. Specifically, researcher might try to represent more realistically the lags between changes in target variables and the policymaker’s perceptions of the changes. More also needs to be done to enhance the comparability of the specification of policy regimes and shocks for models with different time frequencies.” (Bryant, Hooper and Mann 1993, pp. 36) In order to incorporate these lags into committee decision-making or team decision-making, we need to know the properties of each data-set, for example, inflation, GDP, and unemployment, or NAIRU (Non-Accelerating Inflation Rate of Unemployment). This is because, as shown above, optimal decision-making could be based on which information or more specifically, which economic indicators each agent used. Typically, data-generating procedure is very different from the institution-to-institution, indicator-to-indicator, or frequency. Therefore it is very unrealistic for decision-makers to pretend that they make decisions about optimal monetary policy based on current-realization or the revised data of economic indicators. McCallum (1999) argued that ”it is reasonable to assume that contemporaneous observations are available for interest rates, exchange rates, or other asset-market

---

1 The model example for external lag is financial accelerator model. Bernanke, Gertler and Gilchrist 2000. In this model, due to financial market, Federal Funds rate can affect the economy through the interest rates in financial market. Therefore, its effectiveness depends on not only indirect effective rather than direct one but also relatively slower than one without financial sector.
prices. It would be unreasonable, however, to make such an assumption for nominal GDP (or GNP) or price level.” (pp. 1517)

Under a hierarchy structure, therefore, the information updated to agents depends on at which stage the agent is. In other words, most newly updated information is only available to relatively highly ranked decision-maker(s) or final decision-maker(s). In figure 2.1, after period T-1, since newly generated data are only available to final decision-maker due to the lag of data-generating process, the recognition between member 1 and 2 and final decision-maker is different. Therefore since this difference is due to the time difference of access to data-set, its recognition difference could be the one pointed out by Orphanides (2001) and Orphanides (2003). In addition to this, there is one kind of administrative lag for making and announcing the final optimal decision in figure 2.1 since all decisions are not made simultaneously in decision-making procedure.

Moreover, because there is no communication between member 1 and member 2 due to the hierarchical structure in the artificial FOMC, both member 1 and member 2 have different points of view. (Meade and Sheets 2005) This difference is fully based on recognition difference or, as mentioned above, their regional difference. In other words, the regional economic concern which belongs to member 1 or 2 is more important than national-wide economic concern when each Fed committee member decides its monetary policy and this makes each member different even if the same information structure or the same accuracy were available to each member. This implies that they could have regional biased perception, e.g., biased observation which deviates from consensus. Under different data-generating process for each agent due to the hierarchical structure, it is reasonable to regard committee decision-making as a gradual decision process. This is because their data-

---

2McCallum (1999) also mentioned that it is also important to decide how to measure the time-series data, for example, growth-rate or growth-level, that is, difference-stationary or trend-stationary. Of course, traditional view is that important macroeconomic variables are difference-stationary, also supported by Nelson and Plosser (1982). However, if it is employed wrong measurement, policy effect would be very different to its intention.

3Typically, data revision or updating would take much more time than assumed in here. Therefore, it would not be reasonable to make the different access to data. As mentioned above, however, it would be reasonable in case of interest rate, exchange rates, or other asset-market price. This is because the frequency of this data is highly frequent.
generating process heavily relies on the timeliness of the data. In line with this, the decision-making in monetary policy also follows the data-generating process. Therefore gradualism in decision-making is the natural result. (Goodfriend 1991)

Due to gradualism or the inertial movement property, most literature searched for the root of this property in monetary policy. (In figure 2.2, it can be seen that there are no abrupt change in the Federal Funds Rates. At the really least, it can be seen that it takes time for the Federal funds rate to changes direction.) Empirically, the estimation of the coefficient of the lagged interest rate is between 0.6 and 0.8 and statistically significant even with different subsample. (Clarida, Gali and Gertler 2000) Rudebusch (2002), however, argued that the significance of the lagged interest rate would reflect serial correlated errors, and showed low predictive power of term structure for future fund rate in terms of $R^2$. In sum, the reason for interest rate inertia is still controversial and an inconclusive issue even though gradualism is common property in each assertion. Theoretically, there are very different reasons to explain gradualism, such as minimizing financial volatility.
In sum, it could be concluded that the gradualism in monetary policy is based on timeliness of data-generating problem which is due to the discrepancy between revised and real time data, and this could be another reason to assume the gradualism under committee decision-making in this chapter.

![Figure 2.2: Movement of Effective Federal Funds Rates](image)

2.2 Numerical Solution to Derive the Optimal Monetary Policy Under Lags

Given the timeliness of the data due to the hierarchy, the optimal monetary policy can be derived as follows.

Step 1. Construct the auxiliary information structure.

\[ \tilde{z}_{1,t} = z_{1,t} = \xi_{1,t} \quad \text{and} \quad \tilde{z}_{2,t} = z_{2,t} = \xi_{2,t} \]
\[
\begin{bmatrix}
\hat{z}_{1,t+1} \\
\hat{z}_{2,t+1} \\
\hat{z}_{3,t+1}
\end{bmatrix} =
\begin{bmatrix}
z_{1,t} \\
z_{2,t} \\
z_{3,t+1}
\end{bmatrix} + Q^{-1} S \xi
\]

Even though we assume partially nested information structure in order to delete redundant information, there will still be the time lags for decision-making due to delayed observation. Due to partially nested information structure, however, the information structure in a dynamic-team is

\[
\hat{z}_{1,t} = \xi_{1,t} \quad \hat{z}_{2,t} = \xi_{2,t} \quad \hat{z}_{3,t+1} = \xi_{3,t+1}
\]

(2.1)

For simplicity, again, it is assumed that

\[
Q = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad S = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad C = 0 \quad Pr(\xi) = N(0, \sigma_{t}^2)
\]

In here, measurement noise or random disturbance, \( \xi \), implies that it is not white noise. Therefore it has the non-constant variance, but no serial correlation\(^4\) The optimal solution of auxiliary problem given the auxiliary information structure is

\[
\begin{bmatrix}
\tilde{u}_{1,t}^* \\
\tilde{u}_{2,t}^* \\
\tilde{u}_{3,t+1}^*
\end{bmatrix} =
\begin{bmatrix}
3/4 & -1/4 & -1/4 \\
-1/4 & 3/4 & -1/4 \\
-1/4 & -1/4 & 3/4
\end{bmatrix}
\begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix}
\]

\[\xi = \begin{bmatrix}
1/4 \\
1/4 \\
1/4
\end{bmatrix}
\]

\[\tilde{u}_{1,t}^* = \frac{1}{4} \xi \quad \tilde{u}_{2,t}^* = \frac{1}{4} \xi \quad \tilde{u}_{3,t+1}^* = \frac{1}{4} \xi
\]

\(^4\)Due to recognition difference or recognition lags, each has different observation. Therefore its measurement error is different to each other. In this point of view, the non-constant variance is assumed
\[\bar{u}^*_1, t = -\frac{1}{4} \bar{z}_{1,t} \quad \bar{u}^*_2, t = -\frac{1}{4} \bar{z}_{2,t} \quad \bar{u}^*_3, t+1 = -\frac{1}{4} \left[ \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right] \bar{z}_{3,t+1} \quad \text{(2.2)}\]

From (2.2), composite control corresponding to the optimal control for the auxiliary problem is

\[\bar{p}^*_1(\xi) = \bar{p}^*_2(\xi) = \bar{p}^*_3(\xi) = -\frac{1}{4} \xi \quad \text{(2.3)}\]

and

\[\bar{J}^* = E \left[ \frac{1}{2} \begin{bmatrix} \bar{u}^*_1, t & \bar{u}^*_2, t & \bar{u}^*_3, t+1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} \bar{u}^*_1, t \\ \bar{u}^*_2, t \\ \bar{u}^*_3, t+1 \end{bmatrix} + \begin{bmatrix} \bar{u}^*_1, t & \bar{u}^*_2, t & \bar{u}^*_3, t+1 \end{bmatrix} \begin{bmatrix} 1 \\ \xi \end{bmatrix} \right]
\]

\[= E \left[ \frac{1}{2} \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} & -\frac{1}{4} \bar{z}_{2,t} & -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} \\ -\frac{1}{4} \bar{z}_{2,t} \\ -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} + \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} \\ -\frac{1}{4} \bar{z}_{2,t} \\ -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} 1 \\ \xi \end{bmatrix} \right]
\]

\[= E \left[ \frac{1}{2} \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} & -\frac{1}{4} \bar{z}_{2,t} & -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} \\ -\frac{1}{4} \bar{z}_{2,t} \\ -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{1}{4} \bar{z}_{1,t} \\ -\frac{1}{4} \bar{z}_{2,t} \end{bmatrix} \begin{bmatrix} -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \right]
\]

\[+ E \left[ \begin{bmatrix} -\frac{1}{4} \bar{z}_{1,t} & -\frac{1}{4} \bar{z}_{2,t} & -\frac{1}{4} \bar{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} 1 \\ \xi \end{bmatrix} \right]
\]

\[= \frac{1}{2} E \left[ \left( -\frac{1}{4} \bar{z}_{1,t} - \frac{1}{4} \bar{z}_{2,t} - \frac{1}{4} \bar{z}_{3,t+1} \right) \cdot \left( -\frac{1}{4} \bar{z}_{1,t} + \frac{1}{4} \bar{z}_{2,t} - \frac{1}{4} \bar{z}_{3,t+1} \right) \right] + E \left[ \left( -\frac{1}{4} \bar{z}_{1,t} - \frac{1}{4} \bar{z}_{2,t} - \frac{1}{4} \bar{z}_{3,t+1} \right) \cdot \left( -\frac{1}{4} \bar{z}_{3,t+1} \right) \right] + \left. \middle( -\frac{1}{4} \bar{z}_{1,t} - \frac{1}{4} \bar{z}_{2,t} - \frac{1}{4} \bar{z}_{3,t+1} \right) \right] + E \left[ \left( -\frac{1}{4} \bar{z}_{1,t} - \frac{1}{4} \bar{z}_{2,t} - \frac{1}{4} \bar{z}_{3,t+1} \right) \right] \]
\[
= E[-\frac{1}{8} \tilde{z}_{1,t} - \frac{1}{8} \tilde{z}_{2,t} - \frac{1}{8} \tilde{z}_{3,t+1} + \frac{1}{8} \tilde{z}_{1,t} \tilde{z}_{2,t} + \frac{1}{8} \tilde{z}_{1,t} \tilde{z}_{3,t+1} + \frac{1}{8} \tilde{z}_{2,t} \tilde{z}_{3,t+1}]
\]
\[
= -\frac{1}{8} \text{Var}(\tilde{z}_{1,t}) - \frac{1}{8} \text{Var}(\tilde{z}_{2,t}) - \frac{1}{8} \text{Var}(\tilde{z}_{3,t+1}) + \frac{1}{8} \text{Cov}(\tilde{z}_{1,t} \tilde{z}_{2,t}) + \frac{1}{8} \text{Cov}(\tilde{z}_{1,t} \tilde{z}_{3,t+1}) + \frac{1}{8} \text{Cov}(\tilde{z}_{2,t} \tilde{z}_{3,t+1})
\]

Step 2. Substitute \(p^*_i(\xi)\) into \(z_i\)

\[g_1(\xi) = z_{1,t} = \xi\]
\[g_2(\xi) = z_{2,t} = \xi\]
\[g_3(\xi) = \xi + \tilde{p}_1^*(\xi) + \tilde{p}_2^*(\xi) = \frac{1}{2} \xi\]

By corollary, again, if \(g_i\) is invertible for all \(i\), the optimal control laws for the team can be found, and they are equivalent to the optimal control of auxiliary problem, the optimal controls for team are

\[u^*_1,t = \gamma^*(z_{1,t}) = \tilde{p}_1^* g_1^{-1}(z_{1,t}) = -\frac{1}{4} z_{1,t}\]
\[u^*_2,t = \gamma^*(z_{2,t}) = \tilde{p}_2^* g_2^{-1}(z_{2,t}) = -\frac{1}{4} z_{2,t}\]
\[u^*_{3,t+1} = \gamma^*(z_{3,t+1}) = \tilde{p}_3^* g_3^{-1}(z_{3,t+1}) = -\frac{1}{4} \left(\frac{1}{2}\right) z_{3,t+1} = -\frac{1}{2} z_{3,t+1}\]

and the optimal payoff function \(J^*\) corresponding to the optimal control is

\[
J^* = E \left[ \frac{1}{2} \begin{bmatrix} \frac{1}{2} & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{4} z_{1,t} \\ -\frac{1}{2} \tilde{z}_{2,t} \\ -\frac{1}{2} \tilde{z}_{3,t+1} \end{bmatrix} + \begin{bmatrix} -\frac{1}{4} z_{1,t} \\ -\frac{1}{2} \tilde{z}_{2,t} \\ -\frac{1}{2} \tilde{z}_{3,t+1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right] \xi
\]
\[\begin{align*}
= E[-\frac{1}{8} \tilde{z}_1^2, t + \frac{1}{8} \tilde{z}_1, t \tilde{z}_2, t + \frac{1}{4} \tilde{z}_1, t \tilde{z}_3, t+1 + \frac{1}{4} \tilde{z}_2, t \tilde{z}_3, t+1] \\
= -\frac{1}{8} \text{Var}(\tilde{z}_1, t) - \frac{1}{8} \text{Var}(\tilde{z}_2, t) + \frac{1}{8} \text{Cov}(\tilde{z}_1, t \tilde{z}_2, t) + \frac{1}{4} \text{Cov}(\tilde{z}_1, t \tilde{z}_3, t+1) + \frac{1}{4} \text{Cov}(\tilde{z}_2, t \tilde{z}_3, t+1)
\end{align*}\]

Because both \( \hat{J}^* \) and \( J^* \) do not have the same value given control variables, the general solution of this case is not known. Especially this case is the "FIND LINEAR SUBOPTIMAL DECISION FOR \( \Omega \)." (See figure 1.3). To resolve such a problem, the difference between \( \hat{J}^* \) and \( J^* \) is calculated as follows:

\[ \hat{J}^* - J^* = -\frac{1}{8} \text{Var}(\tilde{z}_3, t+1) - \frac{1}{8} \text{Cov}(\tilde{z}_1, t \tilde{z}_3, t+1) - \frac{1}{8} \text{Cov}(\tilde{z}_2, t \tilde{z}_3, t+1) \]

For both to be equal to each other, we must have

\[ \text{Var}(\tilde{z}_3, t+1) = -\text{Cov}(\tilde{z}_1, t \tilde{z}_3, t+1) - \text{Cov}(\tilde{z}_2, t \tilde{z}_3, t+1) \]

Since variance can not have negative value, one of two components on RHS, at least, must a negative value and be larger than the other one. Typically, the more accurate observation member 1 or member 2 makes, the less variance the final decision-maker has. This implies that covariance between the member and final decision-maker is positive. In this case, as mentioned above, general solution would not be available. If a regional bias exists, however, then, the final decision-maker puts more weight on national-wide issue while the member is biased to regional issues. In this case, it could be possible to have a negative covariance.

For example, suppose that the member is concerned about unemployment in his or her region, whereas the final decision-maker is concerned about the inflation rate. In this example, the difference in their observations or interest leads to a negative relation due to the Phillips curve.

---

5This could be one example mentioned in Witsenhausen (1968).
6This argument is based on Meade and Sheets (2005)’s findings.
Because both $\tilde{J}^*$ and $J^*$ could have the same value under this situation, the optimal decision in this team situation is

$$u_{3,t+1}^* = \frac{1}{2} z_{3,t+1} = \frac{1}{2} [u_{1,t}^* + u_{2,t}^* + \xi_{t+1}] = \frac{1}{2} [-\frac{1}{4} \xi_{1,t} + \xi_{3,t} + \xi_{3,t+1}]$$

(2.4)

This equation is rewritten by

$$u_{3,t}^* = -\frac{1}{8} \xi_{1,t-1} - \frac{1}{8} \xi_{2,t-1} + \frac{1}{2} \xi_{3,t}$$

(2.5)

Since optimal decision is function of the state of nature, equation (2.5) gives

$$u_{3,t}^* = -\frac{1}{8} [(a_{1,1} + a_{1,2})(y_{t-1} - y^*) + (a_{2,1} + a_{2,2})(\pi_{t-1} - \pi^*) + \nu_{t-1}]$$

$$+ \frac{1}{2} [a_{1,3}(y_t - y^*) + a_{2,3}(\pi_t - \pi^*) + \nu_t]$$

(2.6)

Where $a_{i,t}$ is the coefficient of each agent on each variables. Because $r_t = a_1(y_t - y^*) + a_2(\pi_t - \pi^*)$, finally, inertial interest rate equation is derived.

$$r_t = \lambda r_{t-1} + \beta_y (y_t - y^*) + \beta_\pi (\pi_t - \pi^*) + \epsilon_t$$

(2.7)

Where $\lambda = -\frac{1}{8}, \beta_\pi = \frac{1}{2} a_{2,3}, \beta_y = \frac{1}{2} a_{1,3}$, and $\epsilon_t = \frac{1}{2} \nu_t - \frac{1}{8} \nu_{t-1}$. This equation implies that inertial interest rate is due to both lagged interest rate in optimal policy and serial correlation in the error term. According to English, Nelson and Sack (2003), allowing both lagged interest rate and serial correlated errors result in both play important role in explaining inertial interest rate. However, the relative importance of lagged interest rate and serial correlation is still open question. In this sense, at least, above inertial policy response under the committee decision-making is successful to capture the dynamics of interest rates.
2.3 Conclusion: Gradualism vs. Persistent Shock

"the debate about the source of gradualism is on going and I cannot hope to render a definitive verdict today on the relative merits of these rationales." Bernanke, Ben, May 20th, 2004 Speech.

The interest rate inertia is one central feature in explaining the central bank’s reaction function. To capture this feature, many literatures have introduced the gradualism or persistent shock in the policy reaction function. Theoretically, incorporating the penalty on the difference in interest rate into the loss function have been introduced to derive the inertial policy reaction function. This is because this allows for interest rate smoothing, that is, gradualism in policy reaction. As pointed out by Brainard (1967), the aggressive policy response to the economic situation under uncertainty would not be optimal. Therefore, it has to be cautious to implement the policy. In this sense, Ben Bernanke indicates the gradualism preference of the Fed in his 2004 speech, "My sense, though, is that policymakers’ caution in the face of many forms of uncertainty and their desire to make policy as predictable as possible both contribute to the gradualist behavior we seem to observe in practice.” In this chapter, even without resorting to incorporate ad-hoc interest rate smoothing in loss function, inertial optimal policy is derived from the lags. Especially, inertial policy response under the committee decision-making captures both partial adjustment and serial correlated errors to account for dynamics of interest rate. As mentioned above, however, this conclusion does not answer if partial adjustment or serial correlated errors has a larger impact on the inertia. This is an empirical matter to which we will turn into the next chapter.

7 If discounting factor in loss function is greater, that is $\beta \rightarrow \infty$, this may lead to less concern about the smoothing. Hence the discounting factor would be restrictive to get smoothing behavior. See Dennis (2006)
8 The Gradual monetary policy is also derived from learning process under uncertainty. See Sack (1998-34)
Chapter 3

Empirical Estimation of the Gradualism

3.1 Introduction

One of main property in policy response to the economy is inertial policy response or the gradualism in policy response. To capture this property, many studies in the literature explore the reason for the inertia. Some studies in the literatures theoretically conclude that the gradualism is the result of minimizing financial volatility (Cuikerman 1991), long term-structure (Goodfriend 1991), or political responsibility, (Goodhart 1997). In contrast, the empirical estimation for the inertia gives partial evidence for the gradualism. In a seminal paper, Clarida, Gali and Gertler (2000) estimated the partial adjustment form, \[ i_t = (1 - \rho)\pi_t^* + \rho i_{t-1}, i_t^* = \pi_t + \beta(\pi_t - \pi^*) + \gamma x_t, \] where \( \pi_t \) is percentage change in price level, \( \pi^* \) is target inflation, and \( x_t \) is a measure of the average output gap. The estimation of the coefficient of lagged interest rate (\( \rho \)) is between 0.6 and 0.8 and statistically significant even with in different subsample. Rudebusch (2002), however, argued that the significance of the lagged interest rate would reflect serially correlated errors, and showed low predictive power of term struc-
ture for future fund rate in terms of $R^2$. When partial adjustment form was employed, Rudebusch (2002) estimated serially correlated shock form, $i_t = \beta \bar{\pi}_t + \gamma x_t + \xi_t$, $\xi_t = \delta \xi_{t-1} + \omega_t$, where $\bar{\pi}_t$ is average (four quarters) inflation and showed that $\delta$ was 0.91 and statistically significant. Even with two different conclusions, English, Nelson and Sack (2003) estimate the policy rule allowing both partial adjustments and serially correlated errors, $\Delta i_t = (1 - \lambda) \Delta i_t + (1 - \lambda)(1 - \rho)(i_t - i_{t-1}) + \lambda \rho \Delta i_{t-1} + \epsilon$, and $i_{t-1} = b_0 + b_\pi \bar{\pi}_t + b_y y_t$, where $\lambda$ and $\rho$ are the partial adjustment parameter and the serial correlation parameter. The first term captures the partial movement of the interest rate to the most recent change while remaining term reflects inertia of interest rate changes. English, Nelson and Sack (2003) show that $\lambda = 0.58$ and $\rho = 0.75$ and conclude that the inertial movement is the result of the sum of the lagged interest rate and serially correlated errors.

To disentangle two different sources of inertia, using DSGE model, Carrillo, Feve and Matheron (2007) estimate both the partial adjustment and serially correlated error parameters with the different parameter restrictions. Under the different parameter restrictions, Carrillo, Feve and Matheron (2007) calculate the Impulse Response Functions (IRFs) of output, inflation, wage inflation, money growth, and Federal Funds rate based on both Structural Vector Autoregression and DSGE model and conclude that the dynamics of each variable is better fitted with moderate gradualism and a high degree of serial correlation. ($\rho = 0.2976, \delta = 0.8740$) In line with this, Coibion and Gorodnichenko (2012) employ real time data, Greenbook data, and estimate the equation which allows both the partial adjustment and serially correlated errors. Even with different datasets, Coibion and Gorodnichenko (2012) show that the coefficient of lagged interest rate and the coefficient of serially correlated errors are 0.83 and 0.55 and are statistically significant.

As this brief literature surveys shows, the empirical literature on the determinants of interest rate inertia is still controversial and inconclusive even though the inertia is common feature in each of these studies. In this chapter, using partial forward-looking model, we will show in our empirical
study that intertial movement is the results of the sum of the lagged interest rate and serially cor-
related error and the inertia due to the lagged coefficient in monetary policy and the persistence
error is more likely to support inertia relative to the persistence errors only. This result would be

3.2 Basic Model

When policymakers set optimal monetary policy, their decision making is subject to the economic
variables which policymakers observe. Therefore the movement or interaction of the economic
variables could be the policy constraints, and the model capturing the movement or interaction of
the economic variables is the important one to explain how policy makers are subject to. In this point
of view, some studies in the literatures, e.g., Dennis (2006), and Ozlale (2003), employed the model
for policy constraints developed by Rudebusch and Svensson (1999). Since this model is backward-
looking model, Soderstrom, Soderlind and Vrdin (2005) allow the degree of forward-looking model
behavior in this model. In theoretical modeling, however, backward-looking and forward-looking
models are still controversial hypotheses. This is because each model gives very different implications
for stability, determinacy, and uniqueness\footnote{For details, see Cochrane (2011)} In this chapter, since the perception about the state
are the main ingredient in setting the policy rule, forward-looking model is more reasonable.

Since output and inflation are most the important constraints to policy makers, the following output
and inflation dynamics are employed:

\[ y_t = a_0 + a_1 y_{t-1} + a_2 E_{t-1} y_t + a_3 [i_t - \pi_t] + g_t \]

\[ \pi_t = b_0 + b_1 \pi_{t-1} + b_2 E_{t-1} \pi_t + b_3 E_{t-1} \pi_{t+1} + b_4 E_{t-1} \pi_{t+2} + b_5 y_{t-1} + \nu_t \]
Where $y_t$ is quarterly real GDP, $\pi_t$ is quarterly CPI on basis of 2005, and $i_t$ is quarterly federal funds rate. As usual $E_{t-1}$ is expectation operator. The error terms, $g_t$ and $\nu_t$, are demand and supply shock, respectively. Compared to Rudebusch and Svensson (1999)'s empirical backward-looking model, this model is partial forward-looking. The practical problem rises here is how to specify the forward-looking expectation. For this purpose, the Greenbook data is used to generate expectation variables. The Greenbook is produced before each meeting of the Federal Open Market Committee. Using an assumption about monetary policy, the Research staff at the Board of Governors prepares projections about how the economy will fare in the future. These projections are made available to the public after a lag of five years. Since the Greenbook data provides the expected future growth rate of annualized percentage change, we employ the formula, $g_m = [(X_m - X_{m-1})^4 - 1] \cdot 100$, to construct the expected value of each variable, where $g_m$ is the annualized percentage change, $X_m$ is the expected value of each variable, and $X_{m-1}$ is the actual data.

In addition the monetary policy rule is as follows:

$$i_t = f_0 + f_1 E_{t-1} \pi_t + f_2 E_{t-1} y_t + f_3 i_{t-1} + \epsilon_t$$

(3.1)

where, following the literature, it is assumed that shocks have long memory and follow an AR(1) process, $\epsilon_t = \rho \epsilon_{t-1} + \xi_t$, where the value of $\rho$ is to be determined empirically. Finally, full system of equations are as follows:

$$y_t = a_0 + a_1 y_{t-1} + a_2 E_{t-1} y_t + a_3 [i_t - \pi_t] + g_t$$

$^2$This rate is the interest rate controlled by the Fed, which each bank charges to each other for unsecured loan for overnight.

$^3$See http://www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data/
\[ \pi_t = b_0 + b_1 \pi_{t-1} + b_2 E_{t-1} \pi_t + b_3 E_{t-1} \pi_{t+1} + b_4 E_{t-1} \pi_{t+2} + b_5 y_{t-1} + \nu_t \]

\[ i_t = f_0 + f_1 E_{t-1} \pi_t + f_2 E_{t-1} y_t + f_3 i_{t-1} + \epsilon_t \]

In this system of equations, the first equation can be seen as the IS curve while the second equation is the Phillips curve, and the last equation is the monetary policy rule. In the Phillips curve, the equation shows that the inflation can be explained by the combination of the weighted sum of expected future inflation and the lagged inflation and output. In the traditional Phillips curve, the inflation depends on the output gap and the expected inflation term, which implies a cost-push effect on the inflation. In order to capture the persistence of the inflation, however, it includes the lagged inflation term. Moreover, output is used instead of the output gap. As mentioned in many studies in the literature, the measurement error of the output gap is large because of the difficulty of measuring the natural level of output. Due to this problem, Gali and Gertler (1999) employed the marginal cost derived from the unit labor cost. However, in this chapter we employ output since output consists of two real marginal costs, labor and capital, under Cobb-Douglas technology. In case of the policy rule, the lagged interest rate term and the persistent shocks are added to traditional Taylor rule to capture the inertia in monetary policy. Given this system of equations, the main objective in this estimation is to check whether there is persistence in the errors, i.e., whether \( \rho \leq 1 \) and the lagged inflation term is significant.

---

4 Most studies in the literature estimated the monetary policy using single-equation. As pointed out by Rudebusch (2002), the additional information, for instance, term-structure, is needed to disentangle the inertia from partial adjustment or serially correlated shock. In this vein, Rudebusch and Wu (2008) employed Macro-Finance model, which includes two aggregate relationships for output and inflation on the basis of term-structure.

5 According to Euler theorem, \( MP_L \cdot L + MP_K \cdot K = Y \). Since \( MP_L \) and \( MP_K \) are equal to real wage and real interest rate, respectively, \( Y \) can be used as cost.
3.3 Estimating the System

To estimate the system of equations, the IS curve, the Phillips curve and the monetary policy can be rewritten as

\[ Y_t = A_0 + A_1 Y_{t-1} + A_2 E_{t-1} Y_t + A_3 E_{t-1} Y_{t+1} + A_4 E_{t-1} Y_{t+2} + \omega_t \]  

(3.2)

Defining

\[
Y_t = \begin{bmatrix} y_t & \pi_t & i_t \end{bmatrix}, \quad \omega_t = \begin{bmatrix} g_t & \nu_t & \epsilon_t \end{bmatrix}, \quad A_0 = \begin{bmatrix} a_0 \\ b_0 \\ f_0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} a_1 & -a_3 & a_3 \\ b_5 & b_1 & 0 \\ 0 & 0 & f_3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_2 & 0 & 0 \\ 0 & b_2 & 0 \\ f_2 & f_1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_4 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\]

In here,

\[
\begin{bmatrix} g_t \\ \nu_t \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho \end{bmatrix} \cdot \begin{bmatrix} g_{t-1} \\ \nu_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} g_t \\ \nu_t \\ \xi_t \end{bmatrix}
\]

The various methods employed in the literature can be split into Maximum Likelihood Estimation (hereafter MLE) and General Methods of Moments (hereafter GMM). In the econometrics literature, the main difference between them is that MLE is not applicable when the full shape of the
distribution of the data is not known. In addition, MLE requires the rational expectation model can be solved only by observable structure or reduced form while GMM does not. Even with this, however, GMM has the problems as 'over-instrumenting' or 'over-correction.' Since the expected variables are known based on the Greenbook data, the rational expectation structure can be fully known. Therefore, we can employ MLE to estimate with less loss of efficiency compared to GMM.

The joint probability density function (PDF) for the data can be defined as follows:

\[
P(\{Y_t\}_1^T; \theta, \Psi) = P(\{Y_t\}_6^T|\{Y_t\}_1^5; \theta, \Psi) \cdot P(\{Y_t\}_1^5; \theta, \Psi) \tag{3.3}
\]

Where \( \Psi \) is variance-covariance matrix of the disturbance vector, \( \omega \), \( \theta = \{a_0, a_1, a_2, a_3, b_0, b_1, b_2, b_3, b_4, f_0, f_1, f_2, f_3, \rho\} \), and \( T \) is sample size including initial conditions. It is postulated that \( \omega_t|\{y_t\}_1^T \sim N(0, \Psi) \) for all \( t \). From (3.2), therefore, the joint probability density function for \( \{Y_t\}_1^T \) can be defined as follows:

\[
P(\{Y_t\}_1^T; \theta, \Psi) = \left[ \frac{1}{(2\pi)^{\frac{n(T-5)}{2}}} |\Psi^{-1}|^{\frac{(T-5)}{2}} (1 - \rho^2)^{\frac{1}{2}} \exp \sum_{t=6}^{T} (-\frac{1}{2} \omega_t'\Psi^{-1}\omega_t) \right] \tag{3.4}
\]

Where \( n \) is the number of endogenous variables, \( y_t, \pi_t, i_t \). From (3.4), quasi-log-likelihood function can be obtained:

\[
\ln L(\theta, \Psi; \{Y_t\}_1^T) \propto \frac{n(T-5)}{2} \ln(2\pi) - \frac{(T-5)}{2} \ln|\Psi| + \frac{1}{2} n(1-\rho^2) - \frac{1}{2} \sum_{t=6}^{T} (\zeta_t'\Psi^{-1}\zeta_t) \tag{3.5}
\]

\(^6\)The problem of GMM in forward-looking model is well described in Mavroeidis (2003). \(^7\)This empirical methodology is based on Dennis (2006). The differences are forward-looking model and parameters in policy rule.
\[\zeta_t = \begin{bmatrix} g_t & \nu_t & \xi_t \end{bmatrix}.'\] Since Quasi Maximum Likelihood Estimation (QMLE) of \(\Psi\) is
\[\hat{\Psi}(\theta) = \sum_{t=6}^{T} \frac{\hat{\omega}_t \hat{\omega}_t^T}{T-5}\]
and concentrated quasi-log-likelihood function is
\[\ln L_c(\theta, \Psi; \{Y_t\}_1^T) \propto -\frac{n(T-5)}{2} \ln(1+2\pi) - \frac{(T-5)}{2} \ln |\hat{\Psi}(\theta)| + \frac{1}{2} \ln(1-\hat{\rho}^2) \tag{3.6}\]

### 3.4 Estimation Result

In the history the regime of Fed is characterized by three regimes: Borrowed Reserves, Non-borrowed Reserves, and Federal Funds rate regime. Since each had very different policy instrument in the regime, it would be reasonable to exclude Borrowed Reserve and Non-borrowed Reserve to estimate. This is because Federal Funds rate is not a policy instrument in these regime. In other words, the monetary policy was implemented by selling and buying government securities under Non-borrowed Reserves or by the control of M1 or M2 under Total Reserves. Therefore, the sample period begins with 1982Q1 and ends in 2007Q3. In this estimation, the variables of interest are the coefficients on lagged Federal Funds rate and persistent policy shock, \(f_3\) and \(\rho\). If the empirical result supports the inertia in monetary policy, then it can, at least, expect that either \(f_3\) or \(\rho\) is positive. The estimation results are shown in Table 3.1.

All coefficients have the conventional sign except \(f_1\), the coefficient on expected inflation rate. According to the Taylor principle, to stabilize the economy, the Fed not only follows the Taylor rule but also it has to raise the nominal interest rate by more than one percentage when the inflation is increased by one percentage point. This implies that \(f_1 > 1\) or at least \(f_1 > 0\). Under the

---

8 see Bernanke and Mihov (1998)
9 After this, the data is not available. This is because the data are released to the public with a lag of five years.
Table 3.1: Quasi-FIML Estimation Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>IS Curve</th>
<th>Phillips Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pointe est.</td>
<td>SD</td>
</tr>
<tr>
<td>$a_0$</td>
<td>26.18138</td>
<td>39.99877</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.124530</td>
<td>0.133366</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.8777646</td>
<td>0.132408</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-0.353350</td>
<td>1.447792</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Federal Funds Rate</th>
<th>Other Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pointe est.</td>
<td>SD</td>
</tr>
<tr>
<td>$f_0$</td>
<td>4.842922</td>
<td>1.659640</td>
</tr>
<tr>
<td>$f_1$</td>
<td>-0.097283</td>
<td>0.043850</td>
</tr>
<tr>
<td>$f_2$</td>
<td>0.000364</td>
<td>0.000201</td>
</tr>
<tr>
<td>$f_3$</td>
<td>0.792079</td>
<td>0.047912</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.556321</td>
<td>0.118724</td>
</tr>
</tbody>
</table>

forward-looking model, however, the Fed expects the future inflation to be lower when policymaker announces future higher interest rate. According to Rudebusch and Wu (2008), moreover, the increase in underlying inflation factor (or perceived inflation objective) has two effect, one is directly increase the short rate, and the other is indirectly decrease the short rate. Since the indirect effect dominates the direct effect, the increase in inflation factor initially lowers the short rate. Most importantly, Taylor rule is satisfied by observed data, that is, revised data. Therefore, it would be reasonable to have a negative coefficient on expected inflation.

In addition, the coefficients on lagged Federal Funds rate and persistent policy shock are statistically significant, consistent with English, Nelson and Sack (2003). This implies that both gradualism and persistent shocks are important factors describing the movements of the Federal Funds rates. In terms of the absolute value, however, the result is more likely to support the gradualism. In other words, this result supports widely held beliefs, interest rate smoothing, that is, there is no sudden change in interest rate since policy response to the economy heavily relies on the previous interest rate. Also, this result is consistent with Clarida, Gali and Gertler (2000), in which the empirical
estimation of lagged interest rate is in between 0.6 and 0.8. In addition, the persistence in the shock is also supporting the explanation of the inertia in interest rate. (Rudebusch 2002)

Figure 3.1 shows the comparison between actual and fitted value from the model. It is shown that the model fits the actual data. Therefore we can conclude that the model does a good job in capturing the movement of each variable including the movements of Federal Fund rate.

3.5 Estimating the System without the Persistent Shock

According to Rudebusch (2002), persistency in Federal Fund rate would be due to the illusion caused by serially correlated shock. The implication behind this argument is that it is difficult to distinguish the gradualism from serially correlated shock when the parameters in policy rules are freely estimated. In this vein, it would be reasonable to estimate the model without the persistent shock although lagged interest rate and the persistent shock have important roles in describing the
inertial movement in the Federal Funds rate. To isolate the effect of the gradualism in this section, the model will be estimated without the persistence in the error term. If the interest rate inertia is due to the serial correlation, then $f_3$ (the lagged interest rate term) should be insignificant.

To estimate the model without persistent shock, the quasi-log-likelihood is modified as follows:

$$lnL(\theta, \Psi; \{Y_t\}_1^T) \propto -\frac{n(T - 5)}{2} ln(2\pi) - \frac{(T - 5)}{2} ln|\Psi| + \frac{1}{2} ln(1 - \rho^2) - \frac{1}{2} \sum_{t=6}^{T} (\omega_t^\prime \Psi^{-1} \omega_t) \quad (3.7)$$

Therefore, concentrated quasi-log-likelihood is

$$lnL_c(\theta, \Psi; \{Y_t\}_1^T) \propto -\frac{n(T - 4)}{2} ln(1 + 2\pi) - \frac{(T - 4)}{2} ln|\hat{\Psi}(\theta)| \quad (3.8)$$

As shown in Table 3.2, all parameters have the conventional sign as before, and in term of absolute value, there is no large difference between this specification and the previous one. In the policy rule, however, only the coefficient on lagged Federal Funds rate is statistically significant, and supports the gradualism. However, since the estimation without the persistence shock shows less statistical significance than the previous one in terms of t-statistic, it would be reasonable to include the persistent shock in the system. Consistent with this, moreover, the model selection criteria, Akaike Information Criterion (AIC) (10.29 vs. 10.62) and Schwarz Information Criterion (SIC) (10.68 vs. 10.98), also favor the system with the lagged interest term and the persistent shock. Also, this result still supports Clarida, Gali and Gertler (2000), in which the empirical estimation of lagged interest rate is in between 0.6 and 0.8.
### Table 3.2: Quasi-FIML Estimation Results without the Persistent Shock

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>27.57362</td>
<td>36.78396</td>
<td>$b_0$</td>
<td>0.508157</td>
<td>0.335335</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.120881</td>
<td>0.137149</td>
<td>$b_1$</td>
<td>0.144982</td>
<td>0.186338</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.881443</td>
<td>0.136109</td>
<td>$b_2$</td>
<td>0.857655</td>
<td>0.059829</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-0.310721</td>
<td>1.398703</td>
<td>$b_3$</td>
<td>0.463693</td>
<td>0.204341</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.482036</td>
<td>0.287319</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.16E-05</td>
<td>4.36E-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Federal Funds Rate</th>
<th>Other Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Pointe est.</td>
</tr>
<tr>
<td>$f_0$</td>
<td>2.795260</td>
</tr>
<tr>
<td>$f_1$</td>
<td>-0.062742</td>
</tr>
<tr>
<td>$f_2$</td>
<td>0.000259</td>
</tr>
<tr>
<td>$f_3$</td>
<td>0.886011</td>
</tr>
</tbody>
</table>

#### 3.6 Estimating the System without the Lagged Interest Rate.

As mentioned above, Rudebusch (2002) showed that the gradualism could be fully explained by the persistence in the errors, that is, serially correlated errors. As before, to isolate the effect of the gradualism, the model will be estimate without the lagged interest rate. The quasi-log-likelihood for the model without the lagged interest rate is:

$$
\ln L(\theta, \Psi; \{Y_t\}_1^T) \propto -\frac{n(T - 5)}{2} \ln(2\pi) - \frac{(T - 5)}{2} \ln |\Psi| - \frac{1}{2} \sum_{t=6}^{T} (\zeta_t' \Psi^{-1} \zeta_t)
$$

(3.9)

Therefore, the concentrated quasi-log-likelihood is:

$$
\ln L_c(\theta, \Psi; \{Y_t\}_1^T) \propto -\frac{n(T - 4)}{2} \ln(1 + 2\pi) - \frac{(T - 4)}{2} \ln |\hat\Psi(\theta)| + \frac{1}{2} \ln(1 - \rho^2)
$$

(3.10)

As shown in Table 3.3, the sign of each coefficient is same as before. All coefficients have the conventional sign. One important conclusion is that the results indicate strong persistence in the error term. The AR(1) coefficient, $\rho$ is statistically significant and close to 1. Even though this
is not direct evidence supporting the gradualism in interest rate, it could be the indirect evidence supporting the inertial movement in interest rate. As mentioned by Rudebusch (2002), therefore, the gradualism could be the illusion. However, it is not conclusive since the gradualism is one reason to account for the inertial movement in the interest rate. This is why the gradualism is still an open question.

Along with this, two main differences between this model and the previous one is (1) the significance of the interest rate equation is marginally improved and (2) the value of the estimate of \( f_1 \) and \( f_2 \) are larger. According to the model selection criteria (AIC and SIC), however, this model is inferior to the estimation with the lagged interest rate and the persistent shock.

| Table 3.3: Quasi-FIML Estimation Results without the Lagged Interest Rate |
|-------------------------|-------------------------|-------------------------|
| IS Curve | Phillips Curve | IS Curve | Phillips Curve |
| Parameter | Pointe est. | SD | Parameter | Pointe est. | SD |
| \( a_0 \) | 25.78153 | 38.10551 | \( b_0 \) | 0.408983 | 0.366992 |
| \( a_1 \) | 0.116297 | 0.131341 | \( b_1 \) | 0.126444 | 0.213963 |
| \( a_2 \) | 0.885727 | 0.130515 | \( b_2 \) | 0.66476 | 0.076228 |
| \( a_3 \) | -0.374605 | 1.417423 | \( b_3 \) | 0.444290 | 0.199757 |
| \( a_4 \) | \( f_4 \) | -0.450577 | 0.292830 |
| \( a_5 \) | \( f_5 \) | 7.24E-05 | 4.28E-05 |

<table>
<thead>
<tr>
<th>Federal Funds Rate</th>
<th>Other Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Pointe est.</td>
</tr>
<tr>
<td>( f_0 )</td>
<td>24.35651</td>
</tr>
<tr>
<td>( f_1 )</td>
<td>-0.450522</td>
</tr>
<tr>
<td>( f_2 )</td>
<td>0.001517</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.889934</td>
</tr>
</tbody>
</table>
3.7 Comparison with the Seemingly Unrelated Regression Estimation

In this section, to check the robustness of the estimation in previous sections, a methodology, Seemingly Unrelated Regression (hereafter SUR), is employed. The results are summarized in Table 3.4, 3.5, and 3.6. As shown in each table, there is in general no significant difference in point estimate. The main difference is that the SUR estimation gives a slightly higher statistical significance in the estimation. In line with the literature, I present here only quasi-FIML estimation result.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>IS Curve</th>
<th>Phillips Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pointe est.</td>
<td>SD</td>
</tr>
<tr>
<td>$a_0$</td>
<td>27.76560</td>
<td>19.75486</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.121863</td>
<td>0.104831</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.880516</td>
<td>0.103520</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-0.301343</td>
<td>0.873369</td>
</tr>
<tr>
<td>$a_4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_0$</td>
<td>4.852812</td>
<td>1.083383</td>
</tr>
<tr>
<td>$f_1$</td>
<td>-0.097559</td>
<td>0.026723</td>
</tr>
<tr>
<td>$f_2$</td>
<td>0.000365</td>
<td>0.000115</td>
</tr>
<tr>
<td>$f_3$</td>
<td>0.791994</td>
<td>0.033155</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.556929</td>
<td>0.090633</td>
</tr>
</tbody>
</table>

3.8 Business Cycle Sensitivity in the Policy

"My topic today is the challenges confronting monetary policy in what has been an unusually weak recovery from a severe recession. I will discuss the Federal Reserve’s ongoing efforts in these circumstances to speed the U.S. economy’s return to maximum employment in a context of price..."
One of the most important objectives in the Federal Reserve System is the unemployment rate. Of course, some central banks have different priorities in their policy. However, stabilizing the economy is the main purpose of the central bank. Unemployment is, in this point of views, at least implicitly an important ingredient in their decision-making. Even if it would not be possible to include unemployment as important economic indicator in the decision-making explicitly, it implicitly one which affects the decision-making. For example, if unemployment rate is high, central bank tries to stimulate the economy with the policy instruments such as interest rates. Therefore it could be possible to test how the observation of the economic indicators like unemployment can have an effect on the policy decision-making. To check this, the data is divided into 2 sub-samples on the basis of unemployment rate. The criterion for the sub-samples is average 6 percent unemployment rate from 1982Q1 to 2007Q3. If the unemployment rate is higher than this, it can be seen as recession, vice versa. Note that the number of observation in each sub-sample below and above 6 percent is 61 and 41, respectively. As shown below in
Table 3.6: SUR Estimation Results without the Lagged Interest Rate

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS Curve</td>
<td></td>
<td></td>
<td>Phillips Curve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_0$</td>
<td>27.51447</td>
<td>19.75714</td>
<td>$b_0$</td>
<td>0.492541</td>
<td>0.225582</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.114017</td>
<td>0.104928</td>
<td>$b_1$</td>
<td>0.119585</td>
<td>0.136181</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.888220</td>
<td>0.103615</td>
<td>$b_2$</td>
<td>0.855748</td>
<td>0.057725</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-0.317131</td>
<td>0.873450</td>
<td>$b_3$</td>
<td>0.487580</td>
<td>0.189448</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$b_4$</td>
<td>-0.478624</td>
<td>0.227939</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$b_5$</td>
<td>8.21E-05</td>
<td>2.99E-05</td>
</tr>
</tbody>
</table>

Federal Funds Rate

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0$</td>
<td>24.38402</td>
<td>0.882758</td>
</tr>
<tr>
<td>$f_1$</td>
<td>-0.451261</td>
<td>0.027163</td>
</tr>
<tr>
<td>$f_2$</td>
<td>0.001520</td>
<td>0.000126</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.889688</td>
<td>0.036903</td>
</tr>
</tbody>
</table>

Table 3.7 and 3.8, the policy responses to the economy is stronger when the economy is below 6 percent, vice versa. In other words, the magnitude of each coefficients under below 6 percent have the larger value than the situation under above 6. This implies that the Fed would be concerned about the employment, that is, the Fed are less responsive when unemployment is high while the Fed are more responsive when unemployment is low. In addition to this, the parameters for the inertia are different to each situation. Under below 6 percent, the parameters have about the same magnitude in each coefficient while it does not under above 6 percent. From this, it can be inferred that the Fed implements the more discretionary policy when there is unemployment above the average while the Fed follows the consistent policy rule when unemployment rate is below. Therefore even if the policy rule does not include some economic indicators, it would be reasonable to think that the Fed considers other indicators when the policy is made. In this point of view, how the Fed observes the economic situation is very important gradient in the decision-making.
Table 3.7: Quasi-FIML Estimation Results with the Sub-Sample above 6 Percent

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>0.137344</td>
<td>5958.432</td>
<td>$b_0$</td>
<td>2.581506</td>
<td>53.11101</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.522840</td>
<td>10.75509</td>
<td>$b_1$</td>
<td>0.213039</td>
<td>23.93276</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.504287</td>
<td>10.58668</td>
<td>$b_2$</td>
<td>0.278487</td>
<td>4.255887</td>
</tr>
<tr>
<td>$a_3$</td>
<td>-2.273485</td>
<td>260.8166</td>
<td>$b_3$</td>
<td>0.187469</td>
<td>11.05993</td>
</tr>
<tr>
<td>$a_4$</td>
<td></td>
<td></td>
<td>$b_4$</td>
<td>0.205622</td>
<td>23.60499</td>
</tr>
<tr>
<td>$a_5$</td>
<td></td>
<td></td>
<td>$b_5$</td>
<td>0.000895</td>
<td>0.014178</td>
</tr>
</tbody>
</table>

Federal Funds Rate Other Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pointe est.</th>
<th>SD</th>
<th>Log likelihood</th>
<th>AIC</th>
<th>SIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0$</td>
<td>3.618587</td>
<td>14.46107</td>
<td>-399.5720</td>
<td>21.89993</td>
<td>21.48967</td>
</tr>
<tr>
<td>$f_1$</td>
<td>0.127016</td>
<td>0.769729</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_2$</td>
<td>-0.001477</td>
<td>0.005837</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_3$</td>
<td>0.532377</td>
<td>0.542595</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.000000</td>
<td>0.002381</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.9 Conclusion: Empirical Approach

In this chapter, empirical results shows that both the gradualism and persistent shock are important to describe the movement of Federal Funds rate. Especially, the interest inertia is more likely to support the gradualism rather than serially correlated errors, in terms of value of coefficient and statistical significance. However, even though the partial adjustment has a more important role in describing the dynamics of the interest rate, we cannot exclude serially correlated errors as did in Clarida, Gali and Gertler (2000). As the model in Chapter 2 shows the inertial movement in the
interest rate is due to the timeliness of the data defined as the discrepancy between real-time data and revised data. Our empirical result shows that the interest inertia is caused by both the partial adjustment and serially correlated errors. We can, therefore, conclude that the interest rate inertia is due to the timeliness of the data. In this point of view, this paper suggests another theoretical reason why the interest rate moves gradually.
Bibliography


