Subjective Beliefs in Monetary Policy Committees

Mike Mariathasan*†
European University Institute (EUI)

February 15, 2009

Abstract

(PRELIMINARY) This paper develops a theoretical framework to understand the impact of monetary policy committees (MPC) on the economy: In contrast to earlier literature, committee members are assumed to have identical preferences and common information. Instead, heterogeneity is introduced through the assumption of subjective beliefs about the true model of the economy. Effects of unofficial decision procedures are analyzed in different economic environments and the degree of consensus and the dispersion of beliefs are identified as crucial parameters. The trade-off between flexibility and robustness is well captured by the majority requirement of the voting rule. While this trade-off has been well understood earlier, previous studies have frequently attempted to capture it by studying the effect of the MPC's size on its ability to stabilize the economy. This paper instead reveals that analyzing size itself is meaningless unless the dispersion of beliefs and the majority requirements are taken into consideration. In fact, contrary to most of the earlier literature it is shown that basically any correlation between performance and committee size is coincidental.

JEL Codes: E52, E58, E61, D71, D83

Key Words: Central Bank Design, Monetary Policy Committee, Beliefs, Model Uncertainty, Learning

*I am especially thankful to Ramon Marimon, Thomas Sargent, Georg Dürnecker, Alexander Kriwoluzky and seminar participants at the European University Institute.

†For comments and suggestions please contact: Mike Mariathasan, European University Institute, Economics Department, Via della Piazzuola 43, 50133 Florence, Italy; Tel. (39) 328 942 8944; Fax (39) 055 4685 902; Mike.Mariathasan@EUI.eu

1 Introduction

Within approximately a decade the responsibility for the conduct of monetary policy has been ceded from single central bank governors to monetary policy committees (MPC).[1] While this transfer of power constitutes a global trend, the particular design of the committee and its decision-making process vary across countries.[2] Observable differences occur mainly with respect to the size and the (formal) decision rule.

While group decision-making is thoroughly studied in microeconomic frameworks and in social psychology, it is not well understood how macroeconomic models should be adjusted if the representative policy-maker is replaced by a committee. The theory has introduced heterogeneity in committees largely through preferences or information and concludes that large committees are more successful in stabilizing the economy.[3] Moreover, the literature usually over-simplifies the decision mechanism by assuming a simple majority voting rule.

In contrast, this paper provides a simple model that introduces heterogeneity among MPC members through subjective beliefs about the true model of the economy. Beliefs are modelled as subjective probabilities in a framework of model uncertainty and learning similar to Sargent (1999) and Cogley and Sargent (2005); The committee in this paper is non-interactive, placing the focus of the analysis on the effect a MPC has on members' ability to infer from the data. The voting rule of the model, a two-stage decision mechanism, enforces truth revelation and has been analyzed by Riboni and Ruge-Murcia (2008) empirically in the context of MPCs. They show that it is a good description of the (informal) decision mechanism of a wide range of central banks. As far as theory is concerned, the mechanism allows to introduce a parameter for the degree of actual consensus in the committee; This parameter and the dispersion of beliefs turn out to be crucial factors. In addition, the framework allows to think more clearly about the effects of committee size. More specifically, it helps to understand why committee size itself is irrelevant. Moreover, it is shown that the model can explain how empirical studies that account only for observable factors such as size or formal decision rules might pick up what seemingly is a U-shaped relationship between size and welfare detrimental variables like inflation.

Although subjectivity is introduced ad-hoc through heterogeneous priors, justification for this approach can easily be provided: Members of the policy board might have attended graduate schools with different economic traditions, they might be econometricians, applied or theoretical economists or no economists at all and they might have spent time in academia or as economic consultants before they were appointed to the MPC.[4] All of these factors influence the way each committee member reads the commonly known data and thereby affects individual beliefs about how the common goal, namely stabilizing the economy, can be achieved most successfully.

Instead the rationale for the traditional approaches is less obvious: Statistical offices usually generate data that is accessible to all members of a monetary policy board. This common availability renders different information sets an unlikely source of heterogeneity.

[1] See Blinder (2004) and Blinder (2006) for a extensive description of this evolution.

[2] Pollard (2004) reports a sample of 88 central banks in which 79 conduct policy by committee.

[3] Note that, if there is only heterogeneity of information this result is a straight-forward implication of Condorcet's jury theorem.

[4] Note that in the particular case of the European Monetary Union also country specific factors are likely to amplify this subjectivity.

The concept of individual preferences, on the other hand, is often not carefully defined. In this paper policy-makers are benevolent and seek to maximize welfare. Preference heterogeneity in an accurate interpretation would therefore imply deviation from the assumption of benevolence; This however is usually not what models of "heterogenous preferences" aim at. If preference heterogeneity is instead solely interpreted as heterogeneity of objective functions, it in fact has an interpretation that involves subjective beliefs. Unlike the model of Cogley and Sargent (2005) suggests, model uncertainty should in principle also imply differences in the (micro-founded) welfare function, i.e. in the objective function of the benevolent policy-maker. Hence, if preference heterogeneity is carefully defined, it is unlikely to play a role, whereas the preference heterogeneity assumed in some models of MPCs has an equivalent interpretation in line with the assumption of subjective beliefs stressed in this paper.[5]

Although the model is not calibrated to a particular economy and therefore does not generate policy advice for a particular currency area, it produces interesting general results: More specifically, it shows that a properly sized committee, using a qualified median voting rule, can outperform a single Bayesian policy-maker in a stable economic environment, if there is model uncertainty and learning. However, if a structural break occurs, the committee will be slow to detect it and therefore will perform worse than a single Bayesian policy-maker, *unless* it changes the decision rule to a median voting scheme. The same intuition applies to the case of a "crisis", a period of increased volatility. However, in this case changing the decision rule towards a more demanding consensus requirement contributes to stabilizing the economy, because the single Bayesian decision-maker falsely interprets large shocks as structural changes.

These observations have interesting institutional implications in favor of the European Central Bank's (ECB) practice to keep minutes of the MPC meetings under locker and key. Such secrecy allows the board to introduce discretion in the informal decision process without generating additional uncertainty in the economy.

The remainder of the paper is structured as follows. Section 2 provides a brief overview of the small body of literature on MPCs and contrasts my approach to comparable work. Section 3 describes the theoretical framework. Section 4 provides the results and Section 5 concludes.

2 Literature

In terms of content this paper clearly belongs to the small body of literature that studies the effects of monetary policy by committee. Contributions to this literature include experimental and empirical work as well as theoretical studies that treat the problem from both a micro- and a macroeconomic perspective.

While the examination of group decision-making has a long tradition in microeconomics and in the field of social psychology other branches have picked it up only more recently.[6] The literature is substantially driven by Alan Blinder and an extensive non-technical summary can be found in

[5] There is of course room for deviations from benevolence in political economy models. However, this would open a whole new subject and is ignored in this work. Yet, the political economy dimension certainly allows for interesting insights in particular for the case of the European Monetary Union.

[6] See Sibert (2006) for a summary of results from social psychology and their presumed impact on central banking by committee. For a more recent contribution to the microeconomic perspective on MPCs see Riboni and Ruge-

Blinder (2006). A standard reference and an initiating paper for a series of follow-up studies is the experimental work by Blinder and Morgan (2005): They find that committees do not only decide better but sometimes also faster than single policy-makers. While the latter result is not undisputed, the fact that committees conduct better policy is a robust characteristic of the experimental literature.

In a very recent paper Berger and Nitsch (2008) take an empirical perspective to assess the impact of committee size on policy quality. They discover a U-shaped relationship between committee size and inflation and infer from this that the trade-off between flexibility and information usually found in group decision-making can be measured through the size of a MPC. While I can recover the U-shape in my analysis my theoretical framework suggests that it is not so much the size of a committee but the decision rule that drives this trade-off.

To my knowledge very few examples of theoretical papers exist that include MPCs in dynamic macroeconomic models. Svensson (2007) considers a committee with heterogeneous preferences but does not address the performance relative to a single policy-maker explicitly. Gerlach-Kristen (2003) conducts an exercise that is similar to this paper: She introduces heterogeneity through noisy signals about the future output gap. In this framework an intuitive application of Condorcet's jury theorem implies that a committee in which signals are aggregated has a better perception of the exact distribution of the signal and therefore conducts more successful policy. My paper instead takes a step into explaining where noisy signals come from. In addition, I show that it is not clear as a general rule that committees are more succesful and it is even less clear - as implied by her framework - that larger committees do always better than small committees. Another paper that is conceptually very close to this work is Berk and Bierut (2004). They also explore the intuitive link between learning models and MPCs. More specifically, they attribute differences in preferred policies to differences in skills and focus on the effect of learning from interaction among committee members. My approach instead is to shut down learning within the committee and to examine how a committee framework affects inference from the data. What I find particularly interesting is the observation that a committee imposes limits on the possibilities of rational experimentation.

In terms of the tools I use in my analysis this paper draws heavily on the standard framework for the study of model uncertainty. Sargent (1999) and Cogley and Sargent (2005) may be mentioned as seminal references. The learning mechanisms are texbook algorithms as summarized for example in Evans and Honkapohja (2001) or Carceles-Poveda and Giannitsarou (2006).

3 Model

I propose a simple model in which there exist two rival theories of the law of motion of the economy that result in two different ways of estimating a Phillips Curve relationship - the keynesian and the classical way. Committee members are aware of the two theories and solve an optimal control problem choosing a policy to minimize a quadratic loss function. The policy instrument is the inflation rate but control of it is incomplete. While agents in this economy know the two rival reduced form models, they have incomplete knowledge about the exact value of the parameters

Murcia (2006) who analyze a dynamic voting game and the interaction of preference heterogeneity with concern about the future. In a two-member committee they are able to raionalize empirically relevant policy inertia.

and estimate them from the (simulated) data using a standard recursive least squares algorithm. In the commitee, policy-makers differ only with respect to the prior probabilities they attach to the two rival models. Each period they update the weights they attach to the loss predicted by each of the models using Bayesian methods and calculate their preferred policy accordingly. Voting in the committee under the chosen voting mechanism is equivalent to choosing a qualified median policy from the set of proposed actions each period. Besides being the empirically relevant voting mechanism for monetary policy boards (Riboni & Ruge-Murcia (2008)) this voting mechanism also ensures truth revelation and hence allows to abstract from the consideration of strategic behaviour.[7] In the entire paper I only consider committees consisting of an odd number of members to avoid draws.

A major drawback of the use of the two reduced form models is that they are not micro-founded. This is problematic because it implies that they do not allow for a sound derivation of the corresponding welfare criterion and thereby of the objective function of a benevolent MPC and of the assesment of the welfare effects. At this writing I nonetheless continue to work with the reduced form models in order to be able to preserve the link to the corresponding literature. However, I repeat the exercise for different welfare criteria in order to account for robustness of my results. In future versions of this project I plan to also account for the micro-foundation in a more rigorous way.

Following Sargent (1999) I denote the deviation of inflation from target at time t by y_t, the deviation of the unemployment rate from target at time t by U_t and the policy variable at time t by $x_{t|t-1}$. The notation underlines that policy is chosen at any date t conditional on all information available up to time $t-1$. As mentioned before, policy-makers' control of inflation is incomplete. Assuming $\xi_t \sim N\left(0, \sigma_\xi^2\right)$ the choice of the policy variable translates into inflation as follows:

$$y_t = x_{t|t-1} + \xi_t, \tag{1}$$

3.1 Phillips Curve(s)

Following Sargent (1999) in his interpretation of King and Watson (1994) I label the two different ways of running a Phillips Curve regression the keynesian and the classical direction of fit. They only differ in that in the *keynesian case* inflation is regressed on unemployment and various lags while in the *classical case* unemployment is regressed on inflation and various lags:

1. Keynesian Fit:
$$y_{t+1} = \gamma_0 + \gamma_1 U_{t+1} + \gamma_a (L) U_t + \gamma_b (L) y_t + \eta_{t+1} \tag{2}$$

2. Classical Fit:
$$U_{t+1} = \beta_0 + \beta_1 y_{t+1} + \beta_a (L) U_t + \beta_b (L) y_t + \varepsilon_{t+1} \tag{3}$$

Where L is the lag operator, $\theta_a(L)$, $\theta_b(L)$ for $\theta \in \{\gamma, \beta\}$ are lag polynomials and η_{t+1} and ε_{t+1} are independently distributed Gaussian errors: $\eta_{t+1} \sim N\left(0, \sigma_\eta^2\right)$ and $\varepsilon_{t+1} \sim N\left(0, \sigma_\varepsilon^2\right)$.

[7] For a sketch of the proof see the appendix.

3.2 Loss Function

As mentioned earlier, the specification of the welfare criterion is ad-hoc as in the correponding literature. This causes a logical inconsistency in the sense that two different structural economic models in general also imply two different welfare criteria. Woodford (2003, Chapter 6) shows that for example a basic New Keynesian model and a model of preset prices, also referred to as a New Classical model differ in that the latter only considers unexpected inflation to be welfare relevant. To account for the inconsistency I conduct my numerical analysis for two different scenarios: In the first case -similar to Cogley and Sargent (2005)- I simply assign one quadratic loss function of the following type to each committee member:

$$\rho\left(U_t, y_t, x_{t|t-1}\right) = -\left(U_t^2 + \lambda \cdot y_t^2\right) \tag{4}$$

The parameter λ identifies how important the policy-maker considers losses from deviation of inflation from target to be relative to losses from movements in unemployment. Svensson (2007) introduces heterogenous preferences via this parameter and has the committee vote on a representative relative weight.

In a second scenario, I take rough account of potential differences in the welfare criterion and assign one loss function to each model. While the specification is guided by the intuition in Woodford (2003, Chapter 6) it should be stressed that these welfare criteria are nonetheless ad-hoc. In future work I plan to change from the econometric models to different types of DSGE models that allow me to construct a welfare criterion from deep parameters.

The welfare criterion corresponding to the classical model takes the following form:

$$\rho^C\left(U_t, y_t, x_{t|t-1}\right) = -\left(\lambda^C \cdot U_t^2 + \sigma_{\xi,t}^2\right) \tag{5}$$

The variance of the residual of equation (4) captures unexpected inflation. In contrast, the loss function assigned to the keynesian model remains basically unchanged as compared to the first scenario:

$$\rho^K\left(U_t, y_t, x_{t|t-1}\right) = -\left(\lambda^K \cdot U_t^2 + \pi_t^2\right) \tag{6}$$

Given data and information at date t each policy member minimizes in expected terms the infinite sum of discounted future losses, assuming his parameter estimates will not change with new vintages of incoming data. Moreover, each policy-maker has a different perception of the likelihood of each of the two models being true in the next periods and chooses his optimal policy such that the weighted average of future expected losses predicted by the rival theories is minimized. Note that incentive compatibility of the voting mechanism ensures that the policy-maker does not need to take into account the fact that not his optimal policy but the choice of the committee will be implemented.

$$\mathcal{L}_i = E_t \sum_{j=0}^{\infty} \delta^j \cdot \left[w_t^i \cdot \rho^C \left(U_{t+j}^C, y_{t+j}^C \right) + \left(1 - w_t^i \right) \cdot \rho^K \left(U_{t+j}^K, y_{t+j}^K \right) \right] \tag{7}$$

Obviously, for the first scenario $\rho^C = \rho^K$. However, it is important to note that in both cases: $\left(U_{t+j}^C, y_{t+j}^C \right) = \left(U_{t+j}^K, y_{t+j}^K \right)$ for $j \leq 0$, but $\left(U_{t+j}^C, y_{t+j}^C \right) \neq \left(U_{t+j}^K, y_{t+j}^K \right)$ $\forall$ $j > 0$.

3.3 Voting

Riboni & Ruge-Murcia (2008) suggest and support empirically a two stage voting mechanism in which committe members decide in a first stage on the direction of change of the policy variable (e.g. to decrease or increase the interest rate) while the decision on the exact increment is found through a super majority vote on successive ϵ-changes in the direction determined during the first stage. While the empirical support is certainly a major virtue of this two-stage consensus mechanism a close inspection reveals that the outcome differs from a qualified median rule by not more than ϵ. Since the choice of the ϵ increment can be made arbitrarily small the outcomes of the two mechanism are effectively equivalent. To illustrate this assume w.l.o.g. an odd number N for the size of the committee and:

$$x_{t+1|t}(1) < y_t < ... < x_{t+1|t}^L < ... < x_{t+1|t}^H < ... < x_{t+1|t}(N) \tag{8}$$

$$x_{t+1|t}^L \equiv x_{t+1|t}\left(\frac{N+1}{2} - S \right); \qquad x_{t+1|t}^H \equiv x_{t+1|t}\left(\frac{N+1}{2} + S \right)$$

In this case, simple majority voting would certainly favour an increase of inflation and successive addition of ϵ to y_t would find the required support of the super majority ($= N - \left(\frac{N+1}{2} - S\right) + 1 = \frac{N+1}{2} + S$) until for some $k \in \mathbb{N}$:

$$y_t + k\epsilon < x_{t+1|t}^L \text{ and } x_{t+1|t}^L < y_t + (k+1)\,\epsilon$$

Hence, $\exists \epsilon^* \in \mathbb{R}$ s.t. for sufficient iterations there exists a large enough k^* for which $y_t + k^*\epsilon^* = x_{t+1|t}^L$. The same is true respectively if $y_t > x_{t+1|t}\left(\frac{N+1}{2} + S\right)$. Hence, for a sufficiently small ϵ the two stage consensus mechanism is identical to a qualified median rule that is characterized by:

$$x_{t+1|t}^* = \begin{cases} x_{t+1|t}^L & \text{, if} \quad y_t < x_{t+1|t}^L \\ y_t & \text{, if} \quad y_t \in \left[x_{t+1|t}^L ; x_{t+1|t}^H \right] \\ x_{t+1|t}^H & \text{, if} \quad y_t > x_{t+1|t}^H \end{cases} \tag{9}$$

The numerical analysis will reveal that the choice of the parameter S is an important one and more importantly a one ignored in both the empirical and the theoretical literature. Although Riboni and Ruge-Murcia (2008) introduce the mechanism to the study of MPCs, they treat the degree of consensus as a parameter rather negligent. However, intuition suggests that the trade-off of committee vs. single-person decision-making is a one between information and flexibility. If this is true the degree of consensus is bound to play a crucial role because it affects most directly exactly this trade-off. As Riboni and Ruge-Murcia (2008) point out however this parameter captures an inofficial decision-making rule; This means that a MPC that officialy operates according to a simple majority vote and a MPC that is dominated by the chairman according to its constitution might effectively decide according to a similar consensus rule simply because the chairman is unwilling to decide against the board or because the formal vote is preceeded by a discussion that generates consensus. Because it is an informal mechanism it is difficult to account for in epirical studies. In particular I show that a U-shaped relationship between size and performance as found by Berger and Nitsch (2008) in a cross-country study does not necessarily imply the existence such a trade-off if the dimension of actual consensus is ignored.

3.4 Updating

3.4.1 Parameter Estimates:

Despite being aware of the two rival models, agents in this economy are not supposed to know the exact value of the parameters within the model. Instead they have to deduct them from their available information. Information at time t consists of observations (U_s, y_s) with $s \in [0, ..., t]$. As is perfectly standard in the macro learning literature agents learn about the parameter in their perceived models by recursively running optimal least squares regressions. As in Evans & Honkapohja (2001) or Carceles-Poveda & Giannitsarou (2006) the updating can be characterized by the following two equations:

$$R_t^n = R_{t-1}^n + g \cdot \left(X_{t-1}^n X_{t-1}^{n\prime} - R_{t-1}^n \right) \tag{10}$$

$$\widehat{\theta}_t = \widehat{\theta}_{t-1} + g \cdot \left(R_t^n \right)^{-1} \cdot X_{t-1}^n \cdot \left(Y_t - X_{t-1}^{n\prime} \widehat{\theta}_{t-1} \right) \tag{11}$$

$$n \in \{C, K\} ; \theta \in \{\beta, \gamma\} ; Y_t \in \{U_t, y_t\}$$

Note that in principle the gain parameter, g, could also vary across models. However, I can think of no consistent argument that would lead a policy-maker to attach more or less weight to more recent observations depending on which model he believes to be more likely. Note in addition, that with a constant g only the distribution of estimates converges but not the estimate itself. With $g = \frac{1}{t-1}$ instead, the algorithm converges and basically corresponds to standard OLS.

3.4.2 Updating of Beliefs:

I assume that agents use Bayesian updating to adjust their beliefs to incoming data. For any given time t the weight committee member i attaches to the loss caused by the classical perception is calculated as follows:

$$w_{t+1}^i = \frac{p\left(U_t|C\right) \cdot w_t^i}{p\left(U_t|C\right) \cdot w_t^i + p\left(U_t|K\right) \cdot \left(1 - w_t^i\right)} \tag{12}$$

Here $p(U_t|n)$ is the conditional likelihood that the realizations observed at time t are generated by model $n \in \{C, K\}$. Using the distributional assumptions on ε_t and η_t this likelihood can be estimated as:

$$p(U_t|n) = \frac{1}{\sqrt{2 \cdot \pi \cdot \widehat{\sigma}^2_{n,t}}} \cdot \exp\left[-\frac{\left(U_t - X_t \cdot \widehat{\theta}_t\right)^2}{2\widehat{\sigma}^2_{n,t}}\right] \tag{13}$$

It is important to be aware of the fact that the Bayesian algorithm implies that once an agent is entirely convinced of one of the two potential models (i.e. $w^i_t = 0$ or $w^i_t = 1$) his beliefs are arrested at this view and no incoming data will cause him to reconsider.

4 Results

4.1 Theoretical Implications of the Framework

4.1.1 The Dynamics of Beliefs:

To examine the dynamics of the weights policy-makers attach to the different models, I start from equation (12). To simplify, I rewrite the expression and define $\omega^i_t \equiv \log\left(\frac{1}{w^i_t} - 1\right)$. Note that $\omega^i_t \in (-\infty, \infty)$ and that ω^i_t is inversely related to w^i_t. It follows:

$$\omega^i_{t+1} = \omega^i_t + \log\frac{p(U_t|K)}{p(U_t|C)} = \omega^i_0 + \sum_{j=0}^{t} \log\frac{p(U_{t-j}|K)}{p(U_{t-j}|C)}; \forall t \tag{14}$$

This already confirms the intuition that policy-makers at each point in time attach more weight to the classical view if they have a higher initial trust in this perception. Moreover they trust more in the classical perception the higher at any point in time the *relative likelihood* of an observation coming from the classical perception as opposed to resulting from a keynesian specification. Notably, equation (14) also implies that the time paths of committee members' beliefs never cross. Hence, once there exists an inactivity region in the aggregation mechanism, i.e. once $L \neq H$, committee decisions will always correspond to the individual decisions of the same two committee members who where ranked at positions L and H. Which of the two is chosen in each period will depend on the desired direction of the policy action. Note that with $L = H$ in the case of simple majority voting this implies that a committee member who initially holds the median belief will continue to hold the median belief over the entire simulation period. Moreover, this committee member will determine the policy of the committee at any point in time. This observation has the important implication that size of the committee alone can not matter. More specifically, adding an arbitrary number of additional committee members does not affect the policy at all if their beliefs are distributed equally below and above the median. Hence, if the degree of consensus is not accounted for discussing committee size alone is meaningless.

In a next step I closer examine the parameters that affect the relative likelihood. For this matter, it is instructive to rewrite the relative likelihood function:

$$\frac{p\left(U_{t-j}|K\right)}{p\left(U_{t-j}|C\right)}=\frac{\widehat{\varepsilon}_{t-j}}{\widehat{\widehat{\eta}}_{t-j}}\cdot\sqrt{\frac{\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\varepsilon}_s}{\widehat{\varepsilon}_{t-j}}\right)^2}{\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\widehat{\eta}}_s}{\widehat{\widehat{\eta}}_{t-j}}\right)^2}}\cdot\exp\left[\frac{[t-j]\left(\left[\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\varepsilon}_s}{\widehat{\varepsilon}_{t-j}}\right)^2\right]^{-1}-\left[\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\widehat{\eta}}_s}{\widehat{\widehat{\eta}}_{t-j}}\right)^2\right]^{-1}\right)}{2}\right];\quad(15)$$

$$\frac{\partial\left[\frac{p(U_{t-j}|K)}{p(U_{t-j}|C)}\right]}{\partial\left[\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\widehat{\eta}}_s}{\widehat{\widehat{\eta}}_{t-j}}\right)^2\right]}>0;\ \frac{\partial\left[\frac{p(U_{t-j}|K)}{p(U_{t-j}|C)}\right]}{\partial\left[\sum\limits_{s=0}^{t-j}\left(\frac{\widehat{\varepsilon}_s}{\widehat{\varepsilon}_{t-j}}\right)^2\right]}<0.$$

Hence, at any point in time the relative likelihood of the latest observation being generated by the keynesian rather than the classical model will be larger the larger the sum of normalized squared residuals generated by the keynesian model over the whole simulation period. The normalization here is with respect to the latest residual. Intuition can be gained when it is realized that the sum will be larger the more frequently the current residual is smaller than any of the past residuals. Ignoring the stochastic influence for the clarity of the argument this means that the current parameter estimates for this model are better than previous ones which obviously will increase trust in this model. Similarly intuitive is the observation that this effect is stronger, the smaller the current residual and the larger this normalized sum relative to the sum generated by the rival model.

In summary, I have identified the prior weight and the sum of squared residuals generated by the two rival models as determining components of the evolution of policy-makers beliefs. However, these residuals depend on the implemented policy which in turn depends on beliefs and the quality of the parameter estimates. Because it is not possible to solve analytically for the policy function, residuals can not be examined in more depth and further analysis is consigned to the quantitative section.

4.1.2 Asymptotic Behavior of the Parameter Estimates

Recursive Least Squares In this section I study the asymptotic behaviour of the parameter estimates if agents employ a recursive least squares learning algorithm. One implication of this mechanism is that unlike in the analysis of the constant gain algorithm in the subsequent paragraph, parameter estimates and therefore the policy function will eventually converge to equilibrium values. I denote these equilibrium values by $U_\infty, y_\infty, \widehat{\gamma}_\infty$ and $\widehat{\beta}_\infty$ respectively. It follows from equations (10) and (11):

$$R_\infty^K=X_\infty^K X_\infty^{K\prime}\Rightarrow X_\infty^K X_\infty^{K\prime}\cdot\widehat{\gamma}_\infty=X_\infty^K y_\infty$$

Partitioning the matrices one arrives at:

$$\begin{bmatrix} U_\infty^2 & U_\infty X'_\infty \\ X_\infty U_\infty & X_\infty X'_\infty \end{bmatrix} \begin{bmatrix} \widehat{\gamma}_{\infty,1} \\ \widehat{\gamma}_{\infty,-1} \end{bmatrix} = \begin{bmatrix} U_\infty \\ X_\infty \end{bmatrix} y_\infty \text{ and } \begin{bmatrix} y_\infty^2 & y_\infty X'_\infty \\ X_\infty y_\infty & X_\infty X'_\infty \end{bmatrix} \begin{bmatrix} \widehat{\beta}_{\infty,1} \\ \widehat{\beta}_{\infty,-1} \end{bmatrix} = \begin{bmatrix} y_\infty \\ X_\infty \end{bmatrix} U_\infty.$$

Having defined the necessary variables I state the following Proposition.

Proposition 1 $\frac{1}{\widehat{\gamma}_{\infty,1}} = \widehat{\beta}_{\infty,1} \Leftrightarrow -\frac{\widehat{\gamma}_{\infty,-1}}{\widehat{\gamma}_{\infty,1}} = \widehat{\beta}_{\infty,-1}$.

Hence, to test for observational equivalence as given by given $(e \cdot \widehat{\gamma}_\infty)^{-1} \cdot (e' - \widehat{\gamma}_\infty) + e' = \widehat{\beta}_\infty$ it remains to be examined if $\frac{1}{\widehat{\gamma}_{\infty,1}} = \widehat{\beta}_{\infty,1}$.[8] Proposition 2 identifies the necessary conditions.

Proposition 2 $(e \cdot \widehat{\gamma}_\infty)^{-1} \cdot (e' - \widehat{\gamma}_\infty) + e' = \widehat{\beta}_\infty \Leftrightarrow \lim_{T\to\infty} F_T^K(i) = \lim_{T\to\infty} F_T^C(i)$,

where:

1. $F_{T,i}^K \equiv (1-g)^{i-1} \cdot \sum_{s=1}^{T} (1-g)^{s-1} \cdot \frac{U_{T-s}}{U_{T-i}} y_{T-s}$
2. $F_{T,i}^C \equiv (1-g)^{i-1} \cdot \sum_{s=1}^{T} (1-g)^{s-1} \cdot \frac{y_{T-s}}{y_{T-i}} y_{T-s}$.

Hence, if $\lim_{T\to\infty} F_T^K(i) = \lim_{T\to\infty} F_T^C(i)$ the two models were observationally equivalent in the long run. It is easy to see however that this can in principle only be expected to hold if $\frac{U_{T-s}}{U_{T-i}}$ and $\frac{y_{T-s}}{y_{T-i}}$ $\forall T, s, i$.[9] Even without stochastic components it is seen immediately that this will not be the case if policy-makers have to learn their parameters because over the learning period the policy function and therefore the value for inflation and unemployment will change with the parameter estimates. What can be learned from the derivation above however, is that there will exist a constant K such that: $\lim_{T\to\infty} F_T^K(i) = K \cdot \lim_{T\to\infty} F_T^C(i)$. This is true, because estimates in this framework will converge in finite time. From this period on the two sums will not continue to grow different and the constant K will capture the differences generated over the transition period. Figure (A) illustrates this finding by plotting the evolution of the estimates of the long run trade off.[10]

[8] Where e is a vector of zeros that has as many rows as γ has columns and as many columns as γ has rows. In addition it features a one to pick up the coefficient on contemporaneous Unemployment.

[9] It might by coincidence also be true if summations cancel out period-wise differences but this is does not constitute a general case I am interested in.

[10] The constant of the long run trade off consists of the constant divided by one minus the sum of all coefficients on unemployment. The slope consists of the sum of all coefficients on inflation divided again by one less the sum of all coefficients on unemployment.

Constant Gain I now turn attention to the case of a constant gain learning algorithm. It is well understood that parameter estimates (and therefore variables) can not be expected to converge to a point.[11] Nonetheless, following Evans & Honkapohja (2001) the asymptotic distribution of the estimates can be approximated as follows:

Proposition 3 *From the distributional assumptions on the error terms in equations (1) and (3) it follows:*

1. $\widehat{\beta}_{\infty} \sim N\left(\beta, g\frac{\sigma_{\varepsilon}^{2}}{2}\left(M_{X^{C}}\right)^{-1}\right)$

2. $\widehat{\gamma}_{\infty} \sim N\left(\left[(e \cdot \beta)^{-1} \cdot (e' - \beta) + e'\right], g\left[\frac{1}{2}\left(M_{X^{K}}\right)^{-1} \widetilde{M}_{X^{K}} \left(M_{X^{K}}\right)^{-1} + \frac{\sigma_{\xi}^{2}}{2}\left(M_{X^{K}}\right)^{-1}\right]\right)$

Where:

- $M_{X^{C}} = E\left[X_t^{C} X_t^{C\prime}\right]$
- $M_{X^{K}} = E\left[X_t^{K} X_t^{K\prime}\right]$
- $\widetilde{M}_{X^{K}} = E\left[X_t^{K} X_t^{K\prime} \cdot \left(PF\left(\widehat{\beta}_t, \widehat{\gamma}_t, \alpha^{i}\right) - \widehat{\gamma}_t\right)\left(PF\left(\widehat{\beta}_t, \widehat{\gamma}_t, \alpha^{i}\right) - \widehat{\gamma}_t\right)' \cdot X_t^{K} X_t^{K\prime}\right]$

Here $PF\left(\widehat{\beta}_t, \widehat{\gamma}_t, \alpha^{i}\right)$ *is the policy function and* X_t^{n} *denote the vectors of right hand side variables in the two rival specifications of the Phillips Curve.*

Proof. The Proof follows very closely Evans & Honkapohja (2001, Section 14.4) who analyze persistent learning dynamics in a smaller version of the model that has been labeled *classical* in this paper. At the core of the proof is the approximation of the stochastic recursive learning algorithm by a continuous-time stochastic differential equation. ■

This section still constitutes work in progress and a full characterization of the M-matrices has not been achieved yet. However, the evolution of the estimates of the long run trade-off, using constant gain estimates in Figure (B) confirm the intuitive result, that additional model uncertainty translates into volatility of the parameter estimates. This should be reflected in the analytical solution for the asymptotic variances.

[11] This is true, because the constant gain will always allow the most recent observation to impact on the estimate. Since the most recent observation always has a stochastic component, estimates can never converge to a point. Note that for this reason convergence could be attained in the deterministic model. In addition, the use of ordinary least squares estimators would shut down the stochastic impact on the estimate in the limit by downweighting recent observations. Values of the variables however would still fluctuate in the limit.

4.1.3 On the Relationship between Committee and Robust decision-making

Given the requirements a committee decision imposes on the likelihood that a particular action is correct, policies conducted by the committee are necessarily more cautious than those of a single Bayesian policy-maker. This raises the question how policy decision-making relates to the most careful way of dealing with model uncertainty: robust decision-making.[12]

Applying the notation of Adam (2004) to the above framework, the problem of a single, robust, policy-maker can be characterized as follows:

$$V_R\left(x_0, S_{-1}\right) = \min_{\{x\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ I_R\left(x_t, S_{t-1}\right) \cdot \rho^C\left(x_t, S_{t-1}\right) + \left(1 - I_R\left(x_t, S_{t-1}\right)\right) \cdot \rho^C\left(x_t, S_{t-1}\right) \right\} \tag{16}$$

Where:

$$I_R\left(x_t, S_{t-1}\right) = \begin{cases} 1 & \text{, if} \quad \rho^C\left(x_t, S_{t-1}\right) > \rho^K\left(x_t, S_{t-1}\right) \\ 0 & \text{, if} \quad \rho^K\left(x_t, S_{t-1}\right) \geq \rho^C\left(x_t, S_{t-1}\right) \end{cases}$$

Similarly, the setup of a Bayesian MPC can be cast into the following form:

$$V_C\left(x_0, S_{-1}\right) = \min_{\{x\}_{t=0}^{\infty}} E_t \sum_{t=0}^{\infty} \beta^t \left\{ \sum_{i=1}^{N} I_C^i \cdot \left[w_i\left(S_{t-1}\right) \cdot \rho^C\left(x_t, S_{t-1}\right) + \left(1 - w_i\left(S_{t-1}\right)\right) \cdot \rho^K\left(x_t, S_{t-1}\right) \right] \right\} \tag{17}$$

Where:

$$I_C^i\left(x_t, S_{t-1}\right) = \begin{cases} 1 & \text{, if} \quad \begin{array}{c} \left[\left(x_{t-1} < x_t^L\right) \wedge (i = L)\right] \\ \vee \left[\left(x_{t-1} \in \left[x_t^L; x_t^H\right]\right) \wedge (i = j)\right] \\ \vee \left[\left(x_{t-1} > x_t^H\right) \wedge (i = H)\right] \end{array} \\ 0 & \text{, if} \quad \begin{array}{c} \left[\left(x_{t-1} < x_t^L\right) \wedge (i \neq L)\right] \\ \vee \left[\left(x_{t-1} \in \left[x_t^L; x_t^H\right]\right) \wedge (i \neq j)\right] \\ \vee \left[\left(x_{t-1} > x_t^H\right) \wedge (i \neq H)\right] \end{array} \end{cases}$$

For simplicity I separate the choice variable x from the state vector S_{-1}. The latter includes the lagged observations on unemployment and inflation but also the currently estimated parameters; (16) and (17) can then be reformulated recursively:

$$V_R\left(x, S_{-1}\right) = \min_x \left\{ \begin{array}{c} \rho^C\left(x, S_{-1}\right) - I_R\left(x, S_{-1}\right) \cdot \left(\rho^K\left(x, S_{-1}\right) - \rho^K\left(x, S_{-1}\right)\right) \\ +\beta \cdot E\left[V_R\left(x', S\right)\right] \end{array} \right\} \tag{18}$$

[12] Following Adam (2004), 'robust decision making' is used synonymously to 'maxmin decision making' as defined by Gilboa and Schmeidler (1989) throughout the paper.

$$V_C(x, S_{-1}) = \min_x \left\{ \begin{array}{c} \mathcal{L}_K(x, S_{-1}) - \left(\sum_{i=1}^{N} I_C^i(x, S_{-1}) w_i(S_{-1}) \right) \cdot (\mathcal{L}_K(x, S_{-1}) - \mathcal{L}_C(x, S_{-1})) \\ +\beta \cdot E[V_C(x', S)] \end{array} \right\} \tag{19}$$

From which it can be seen that:

$$V_R(x, S_{-1}) = V_C(x, S_{-1}) \Leftrightarrow I_R(x, S_{-1}) = \sum_{i=1}^{N} I_C^i(x, S_{-1}) w_i(S_{-1}) \tag{20}$$

Equation (20) explicates two necessary conditions for the equivalence of committee with robust decision-making: Firstly, the set of beliefs represented in the committee must necessarily contain $w_k = 1$ and $w_l = 0$ for at least one k and one $l \in [1, ..., N]$. If this were not true, the right hand side of equation (20) could never exactly equal zero or one.[13] Since, strict belief of this sort in a Bayesian learning environment implies that these beliefs will never be corrected this condition is fairly strong if taken at face value. Even more interesting however is the observation that the committee features an inactive region which a robust decision-maker lacks in principle. Equality of policy decision can still arise in two special cases - one of which will occur with probability zero only while the second one is trivial. Firstly, a robust decision-maker would choose to remain inactive at a fixed point of his policy function. However, in a stochastic model this would necessarily demand that the disturbance term takes on the same value in subsequent periods which occurs with probability zero only. Secondly, robust policy can only be equal to committee policy if the committe never decides to remain inactive. Under the assumed voting scheme this necessarily requires simple majority vote, i.e. a median voter decision: $L = H \ (\equiv M)$. The latter condition simplifies the indicator function for the committee problem:

$$I_C^i = \begin{cases} 1 & , \text{if} \quad i = M \\ 0 & , \text{if} \quad i \neq M \end{cases}$$

Hence, imposing this condition on equation (20) implies:

$$I_R = \sum_{i=1}^{N} I_C^i w_i = w_M \tag{21}$$

Clearly, this requires the median committee member to be a robust decision-maker. Since the time path of beliefs can never cross if a Bayesian learning scheme is assumed this also requires each single committee member to decide according to a robust decision mechanism.

[13] This is true because beliefs can never take on negative values.

The above analysis reveals that the likelihood requirements a committee imposes on a decision limit extreme policies in two ways: On the one hand, an action is only taken if it is sufficiently likely to be correct. On the other hand, an action that minimizes potential future regret is rather not taken if the expected value of the regret relative to the expected loss incured if the state that justifies the safe strategy does not materialize is sufficiently small. In both cases sufficiency is determined by the size of the required super majority, the parameter S that has been pointed out before.

This argument suggests that the benefit of a decision committee results from the fact that it does not fall prey to (potentially) misleading evidence as easy as a single Bayesian decision-maker but at the same time preserves some degree of flexibility by qualifying the 'maxmin decision principle' with the aggregation of subjective beliefs.

4.2 Numerical Analysis

4.2.1 Setup

For the numerical analysis I limit the observation period to 40 observations unless indicated otherwise. The effect of subjective beliefs has materialized during this period and longer horizons would only pick up the persisting differences of the parameter estimates. Once beliefs have converged the difference between a committee and a single decision-maker and thereby the focus of this paper disappears. In addition, I consider 10 years on average a reasonable period for one central bank government to remain in office. The change of parts of the governing board might then be interpreted as the introduction of a fresh set of beliefs.

I always consider the maximum dispersion of beliefs. Hence, when committees of any size are considered the set of beliefs is always equally distributed between ≈ 0 and ≈ 1. The reason not to set the boundaries exactly equal to zero and one is simply to allow every committee member to revise his beliefs.

Throughout the analysis I will consider three different models for the actual law of motion (ALM) for the economy. A classical and a keynesian model of the Phillips curve as considered by the committee and a natural rate version of the Phillips curve that Cogley and Sargent (2005) introduce as the Solow/ Tobin model:

1. Model A: *Classical Phillips Curve*

$$U_{t+1} = \begin{bmatrix} 0.2974 \\ -0.2016 \\ 0.9302 \\ 0.0024 \\ 0.2064 \\ -0.0140 \\ 0.1047 \\ 0.0290 \end{bmatrix}' \begin{bmatrix} 1 \\ y_{t+1} \\ U_t \\ U_{t-1} \\ y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \end{bmatrix} + \varepsilon_{t+1}$$

$$\varepsilon_{t+1} \sim N\left(0; 0.1022\right)$$

2. Model B: *Keynesian Phillips Curve*

$$y_{t+1} = \begin{bmatrix} 2.0943 \\ -0.1754 \\ 0.2304 \\ -0.1933 \\ 0.6589 \\ 0.0589 \\ 0.2623 \\ -0.4106 \end{bmatrix}' \begin{bmatrix} 1 \\ y_{t+1} \\ U_t \\ U_{t-1} \\ y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \end{bmatrix} + \eta_{t+1}$$

$$\eta_{t+1} \sim N(0; 0.0982)$$

3. Model C: *Solow/ Tobin Phillips Curve*

$$u_0^n = U_0$$

$$u_{t+1}^n = u_t^n + 0.075 \cdot (U_t - u_t^n)$$

$$y_{t+1} - y_t = \begin{bmatrix} -0.3923 \\ -0.6871 \\ 0.3040 \\ 0.5291 \\ 0.0074 \\ 0.0325 \end{bmatrix}' \begin{bmatrix} (U_{t+1} - u_{t+1}^n) \\ (U_t - u_t^n) \\ (U_{t-1} - u_{t-1}^n) \\ (y_t - y_{t-1}) \\ (y_{t-1} - y_{t-2}) \\ (y_{t-2} - y_{t-3}) \\ y_{t-2} \\ y_{t-3} \end{bmatrix} + \eta_{t+1}^{ST}$$

$$\eta_{t+1}^{ST} \sim N(0; 0.0619)$$

All values are estimates on quaterly inflation and unemployment data for the Euro Area (1995Q1-2008Q1). However, since this is not an empirical exercise the data mainly serves to place the calculations roughly in the right range.

To initiate the recursive learning algorithms I use parameter estimates on half the sample. These have proven to be convenient starting values. Moreover, it appears to be a reasonable assumption that an incoming MPC has partial information at the beginning of a term. More specifically $R_0^K = R_0^C = I_{8x8}$ and:

$$\widehat{\beta}_0 = \begin{bmatrix} -0.1199 \\ -0.1884 \\ 0.8872 \\ 0.0821 \\ 0.1887 \\ -0.0555 \\ 0.0979 \\ 0.1195 \end{bmatrix}; \widehat{\gamma}_0 = \begin{bmatrix} 1.2578 \\ -0.0994 \\ 0.0738 \\ -0.0650 \\ 0.4071 \\ 0.4529 \\ 0.5966 \\ -0.6809 \end{bmatrix}$$

The remaining parameters are set as follows:

g	β	λ	$\lambda^K = \lambda^C$
0.015	0.99	16	0.04

Except for the relative weights on unemployment these values are standard. The chosen value is in line with the possible ranges for deep parameters a basic New Keynesian model suggests, but since my loss function features unemployment rather then the efficiency gap for output, not much weight should be put on this choice - qualitative results are not dependend on this particular choice.

4.2.2 Baseline Analysis

The setup described in the previous section is now used to evaluate the performance of a committee. To do so, the existence of a committee is taken as given and the relative performance of different combinations of committee and super majority size is evaluated. The exact size of the welfare losses is more or less irrelevant for this exercise since a realistic measure would certainly require a more careful choice of parameters and starting values. What is relevant is the relative performance of different committees within the environment of the three models.[14]

For a first account I study the performance within a stable environment, meaning that the variance of the shocks is kept moderate as described in the description of the ALMs. The results are shown in Figures (C) to (E). While the welfare functions vary with the models the qualitative results remain unaltered also with different specifications of the welfare criterion. For Model A and B the figures show the result under the assumption of different welfare criteria across models, while the results for Model C assume a common criterion. Two observations stand out: First, a full unanimity requirement (observations along the diagonal) seems to limit the committee's ability to respond to the economy too much. Second, the simulations suggest the existence of a U-shape similar to the one found by Berger and Nitsch (2008);[15] However, this shape occurs along the dimension of the super majority. Missing account of this variable can therefore seemingly translate the shape into an apparent relation between size and performance even when -as in the case of simple majority voting and equally distributed beliefs- size has no effect at all.

4.2.3 Disperion of Beliefs

To further highlight the result of the previous section I conduct the analysis for Model B again, this time reducing the range of beliefs among committee members to the intervall $(0.25; 0.75)$. Clearly the shape of the welfare plot (H) changes and, again, suggests that studies with a focus on the size of a MPC miss crucial elements. Interestingly, the trade-off shifts towards a more demanding consensus requirement. This is the case because now the smaller range of beliefs already enhances the flexibility of the committee - hence, given that the committee as a whole is able to respond more quickly it is now more desirable to increase robustness through the voting scheme. This shows that the two factors, the degree of consensus in the voting mechanism and the dispersion of beliefs across committee members are by no means independent.

[14] Welfare measures are always scaled by the maximum loss created by the worst performing committee.

[15] Berger and Nitsch (2008) only observe inflation as a proxy for welfare losses. Again, the qualitative results do not change if I observe for example average inflation only.

What makes these comments relevant for empirical work is that both, informal decision rules within the committee and the dispersion of MPC members' beliefs, are factors that are hard to control for in practice.

Another interesting observation is that the median voting committee result can also be interpreted as a single Bayesian decision-maker. In this case the different scenarios present the robust result that a comittee -if constructed well- can potentially outperfom a single decision-maker. The intuition here is again that the parameter S describing the super majority captures the trade off between flexibility and robustness and according to my analysis some degree of robustness, i.e. a positive S seems to improves this trade off.

In previous versions of the paper I have also evaluated the performance of the committee relativ to single decision-makers that either know the true model or do not revise beliefs in the wrong model. However, these comparisons usually produce exactly the expected result: Knowledge of the true model leads to better performance, strict belief in the wrong model produces the worst outcome. An interesting case however arises in the case of Model C. Since both models A and B can be fit to the data produced by Model C reasonably well it actually turns out to be beneficial to conduct policy according to any of the wrong models.[16] The drawback introduced by the model uncertainty (not necessarily, but also through a committee) is that beliefs shift from one model to the other depending on realizations of the shock and hence policy is sometimes conducted according to one and sometimes according to another model - this additional uncertainty generates uncertainty about the parameter estimates and this volatility turns out to be welfare decreasing. Figure (G) illustrates this alternation of beliefs. Apart from this I abstract from the analysis of strictly believing agents in this version of the paper.

4.2.4 Structural Break

In the previous section the performance of committees has been analyzed in a stable environment and a trade-off between a committee's ability to respond and its robustness to respond too easily to misguiding evidence has been identified. The question arises if and how this trade-off changes if the environment is less stable. For this reason, the case of a structural break is considered. In particular, after 20 quarters the ALM is assumed to switch from Model B to Model A. While this break certainly overemphasizes what would be observed in reality it also serves to clarify the underlying intuition. Figure (F) shows that with the structural break the trade-off shifts in such a way as to make flexibility more valuable. A committee with a simple majority voting rule (or alternatively a single Bayesian decision-maker) outperform committees with a more demanding decision rule simply because it detects the break earlier.

This observation suggests that it would be desirable to endogenize the (informal) decision rule of the MPC. For practical implementation however two difficulties arise: On the one hand endogenizing the decision rule would require knowledge about the economic environment; The committee would be required to know that it is facing a structural break rather than noise. On the other hand a discretionary decision mechanism is likely to jeopardize the achievements of the theory of

[16]Since the Solow/ Tobin model has a Keynesian core it is not surprising that the keynesian Model B can be fit better.

independent and conservative central bankers; That is, while it might improve the interest rate decisions ceteris paribus it is also likely to generate additional uncertainty.

Hence, while practical implications of this finding are unclear the result points out again that earlier studies with a focus mainly on the impact of committee size are likely to miss important determinants. It might simply be the case that a cross country study picks up different economic environments for committees with otherwise identical decision requirements - the particular relation to committee size would then be purely coincidental.[17]

4.2.5 The Case of a Crisis

Finally, a situation of increased economic volatility is analyzed. Contrary to the previous section however, the underlying model of the economy remains unchanged. The ALM is again assumed to be Model B but the variance σ_{η}^{2} is increased from 0.0982 to 4. In Figure (I) it can be seen that the trade-off shifts again towards stronger consensus requirements. The interpretation for this observation is that now the more flexible committee (or alternatively, the single Bayesian decision-maker) falls prey to misguiding evidence too quickly. In particular, it mistakes particular large shocks for evidence for or against a particular model. A more cautious committee instead demands higher likelihood requirements in order to adjust policy to shifting beliefs. If the law of motion of the economy does not actually change but is only harder to detect, these requirements actually contribute to a stable environment.

Clearly this observation underlines the difficulties that arise with respect to practical implementability. However, theoretically one can conclude that unless the economy changes structurally, a super majority demand requirement in a MPC helps to stabilize the economy in particular with respect to a representative Bayesian decision maker or with respect to the simple majority voting rule that is frequently assumed in the theoretical literature. It is beneficial to tighten this requirement in volatile times and it is advisable to loosen it if the economy undergoes a structural change.

5 Conclusion

This analysis introduces a MPC into a dynamic macroeconomic model that features model uncertainty and learning similar to Sargent (1999) and Cogley and Sargent (2005). The analysis directs the attention towards two important factors that have been treated negligent in the existing literature: The dispersion of beliefs across the members of the committee and the degree of consensus that the committee requires to agree on a policy change. These observations have important implications for theoretical and empirical work.

In terms of theory it is pointed out that a median voter assumption simplifies the analysis too much. More specifically, it neglects one important dimension along which a committee adds robustness to its decisions: the degree of consensus. A second theoretical contribution is implied by the observation that committee size alone is not important. This result puts research to the

[17] Note that "different economic environments" does not refer to different positions on the business cycle but to structural differences of the economies.

test that has attempted to characterize the optimal monetary policy committee by its size. Taking the results of this paper at face value, these previous results are purely coincidental. Finally, the numerical analysis shows that the optimal committee can not be a static entity. In particular the informal decision routine should adjust to the economic environment.[18] I interpret this result in favor of the secrecy the ECB employs with respect to the publication of minutes. The discretion, the analysis in the previous sections calls for, is accompanied by the danger of destabilizing market expectations; This danger is mitigated if the degree of consensus is not publicly known.

As far as empirical analysis is concerned this paper again puts previous results to the test. The fact that two, in principle unobservable, factors affect the committee's ability to stabilize the economy crucially requires careful empirical analysis. As Riboni and Ruge-Murcia (2008) point out for example, the formal decision scheme does not provide a sufficient proxy for the informal procedure. It should be interesting to see the effect a good proxy has on the results of earlier empirical work.

This empirical path also points the direction for future research. More specifically, fitting the model to particular currency areas can not only provide more specific advice for the conduct of monetary policy by committee, it will also allow to evaluate the model's predictions relative to historic data across different countries and periods.

[18] This is a theoretical result. The transmission towards practical implementability is referred to future research.

References

[1] Adam, K. 2004. "On the relation between robust and Bayesian decision making." *Journal of Economic Dynamics & Control* (28): 2105-2117.

[2] Berger, H. and V. Nitsch. 2008. "Too many Cooks? Committees in Monetary Policy." *CESifo Working Paper* 2274.

[3] Berk, J.M. and B.K. Bierut. 2004. "The effects of learning in interactive monetary policy committees." *Netherlands Central Bank MEB Series* 2004-01.

[4] Blinder, A.S. and J. Morgan. 2005. "Are Two Heads Better than One? Monetary Policy by Committee." *Journal of Money, Credit and Banking* 37(5): 789-812.

[5] Blinder, A.S. 2006. "Monetary Policy by Committee: Why and How?" *DNB Working Paper* 092.

[6] ———. 2004. *The Quiet Revolution. Central Banking Goes Modern.* New Haven, CT: Yale University Press.

[7] Cogley, T. and T.J. Sargent. 2005. "The Conquest of US Inflation: Learning and Robustness to Model Uncertainty." *Review of Economic Dynamics* 8(2): 528-563.

[8] Evans, G.W. and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics.* Princeton, N.J.: Princeton University Press.

[9] Gerlach-Kristen, P. 2006. "Monetary Policy Committees and Interest Rate Setting." *European Economic Review* 50(2): 487-507.

[10] Gilboa, I. and D. Schmeidler. 1989. "Maxmin expected utility theory with non-unique prior." *Journal of Mathematical Economics* (18): 141–153.

[11] King, R.G. and M.W. Watson. 1994. "The Post-War US Phillips Curve: A Revisionist Econometric History." *Carnegie-Rochester Conference Series on Public Policy* 41(1): 157-219.

[12] Riboni, A. and F.J. Ruge-Murcia. 2006. "The Dynamic (In)efficiency of Monetary Policy by Committee." *Université de Montréal Working Paper 2006-02.*

[13] ———. 2008. "Monetary Policy by Committee: Consensus, Chairman Dominance or Simple Majority?" *Université de Montréal Working Paper* 2008-02.

[14] Sibert, A.C. 2006. "Central Banking by Committee." *CEPR Discussion Paper* 5626.

[15] Sargent, T.J. 1999. *The Conquest of American Inflation.* Princeton, N.J.: Princeton University Press.

[16] Svensson, L.E.O. 2007. "Optimal Inflation Targeting: Further Developments of Inflation Targeting." *Central Bank of Chile Working Paper* 403.

[17] Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, N.J.: Princeton University Press.

Appendix

A Proofs.

Proposition 1. Rewriting the equations from the partition:

$$U_\infty^2 \widehat{\gamma}_{\infty,1} + U_\infty X'_\infty \widehat{\gamma}_{\infty,-1} = U_\infty y_\infty \tag{22}$$

$$X_\infty U_\infty \widehat{\gamma}_{\infty,1} + X_\infty X'_\infty \widehat{\gamma}_{\infty,-1} = X_\infty y_\infty \tag{23}$$

$$y_\infty^2 \widehat{\beta}_{\infty,1} + y_\infty X'_\infty \widehat{\beta}_{\infty,-1} = y_\infty U_\infty \tag{24}$$

$$X_\infty y_\infty \widehat{\beta}_{\infty,1} + X_\infty X'_\infty \widehat{\beta}_{\infty,-1} = X_\infty U_\infty \tag{25}$$

Equation (22) and equation (24) result directly in:

$$\widehat{\gamma}_{\infty,1} = U_\infty^{-1} y_\infty - U_\infty^{-1} X'_\infty \widehat{\gamma}_{\infty,-1}; \qquad \widehat{\beta}_{\infty,1} = y_\infty^{-1} U_\infty - y_\infty^{-1} X'_\infty \widehat{\beta}_{\infty,-1}.$$

This allows to proof the proposition as follows:

$$\frac{1}{\widehat{\gamma}_{\infty,1}} = \widehat{\beta}_{\infty,1} \Leftrightarrow 1 = \left(U_\infty^{-1} y_\infty - U_\infty^{-1} X'_\infty \widehat{\gamma}_{\infty,-1}\right)\left(y_\infty^{-1} U_\infty - y_\infty^{-1} X'_\infty \widehat{\beta}_{\infty,-1}\right)$$
$$\Leftrightarrow$$
$$1 = 1 - U_\infty^{-1} X'_\infty \widehat{\beta}_{\infty,-1} - X'_\infty \widehat{\gamma}_{\infty,-1} y_\infty^{-1} + U_\infty^{-1} X'_\infty \widehat{\gamma}_{\infty,-1} y_\infty^{-1} X'_\infty \widehat{\beta}_{\infty,-1}$$
$$\Leftrightarrow$$
$$U_\infty X'_\infty \widehat{\gamma}_{\infty,-1} = \left(X'_\infty \widehat{\gamma}_{\infty,-1} - y_\infty\right) X'_\infty \widehat{\beta}_{\infty,-1} \Leftrightarrow U_\infty X'_\infty \widehat{\gamma}_{\infty,-1} = \left(-U_\infty \widehat{\gamma}_{\infty,1}\right) X'_\infty \widehat{\beta}_{\infty,-1}$$
$$\Leftrightarrow$$
$$X'_\infty \widehat{\gamma}_{\infty,-1} = \left(-\widehat{\gamma}_{\infty,1}\right) X'_\infty \widehat{\beta}_{\infty,-1} \Leftrightarrow -\frac{\widehat{\gamma}_{\infty,-1}}{\widehat{\gamma}_{\infty,1}} = \widehat{\beta}_{\infty,-1}$$

■

Proposition 2. First, note:

$$\widehat{\theta}_t = \left[\sum_{i=1}^{t} (1-g)^{i-1} \cdot \left(X_{t-i}^n X_{t-i}^{n\prime}\right)\right]^{-1} \left[\sum_{i=1}^{t} (1-g)^{i-1} \cdot \left(X_{t-i}^n Y_{t+1-i}\right)\right]$$

Again, partition the vector of parameter estimates as done before to obtain the following two sets of equations:

$$\begin{bmatrix} A_T^K \cdot \widehat{\gamma}_{T,1} + B_T^K \cdot \widehat{\gamma}_{T,-1} \\ C_T^K \cdot \widehat{\gamma}_{T,1} + D_T^K \cdot \widehat{\gamma}_{T,-1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{T} (1-g)^{i-1} \cdot U_{T-i} y_{T-i} \\ C_T^C \end{bmatrix} \tag{26}$$

$$\begin{bmatrix} A_T^C \cdot \widehat{\beta}_{T,1} + B_T^C \cdot \widehat{\beta}_{T,-1} \\ C_T^C \cdot \widehat{\beta}_{T,1} + D_T^C \cdot \widehat{\beta}_{T,-1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{T} (1-g)^{i-1} \cdot y_{T-i} U_{T-i} \\ C_T^K \end{bmatrix}. \tag{27}$$

Where the following definitions have been used:

$$\left[\begin{array}{cc} A_T^K & B_T^K \\ C_T^K & D_T^K \end{array}\right] = \left[\begin{array}{cc} \sum_{i=1}^{T}(1-g)^{i-1}\cdot U_{T-i}^2 & \sum_{i=1}^{T}(1-g)^{i-1}\cdot U_{T-i}X'_{T-i} \\ \sum_{i=1}^{T}(1-g)^{i-1}\cdot X_{T-i}U_{T-i} & \sum_{i=1}^{T}(1-g)^{i-1}\cdot X_{T-i}X'_{T-i} \end{array}\right];$$

$$\left[\begin{array}{cc} A_T^C & B_T^C \\ C_T^C & D_T^C \end{array}\right] = \left[\begin{array}{cc} \sum_{i=1}^{T}(1-g)^{i-1}\cdot y_{T-i}^2 & \sum_{i=1}^{T}(1-g)^{i-1}\cdot y_{T-i}X'_{T-i} \\ \sum_{i=1}^{T}(1-g)^{i-1}\cdot X_{T-i}y_{T-i} & \sum_{i=1}^{T}(1-g)^{i-1}\cdot X_{T-i}X'_{T-i} \end{array}\right];$$

$$\left[\begin{array}{c} E_T^K \\ E_T^C \end{array}\right] = \left[\begin{array}{c} \sum_{i=1}^{T}(1-g)^{i-1}\cdot U_{T-i}y_{T-i} \\ \sum_{i=1}^{T}(1-g)^{i-1}\cdot y_{T-i}U_{T-i} \end{array}\right].$$

From equations (26) and (27) one obtains the following expressions for $\widehat{\gamma}_{T,1}$ and $\widehat{\beta}_{T,1}$:

$$\widehat{\gamma}_{T,1} = \left[A_T^K - B_T^K\cdot\left(D_T^K\right)^{-1}\cdot C_T^K\right]^{-1}\left[E_T^K - B_T^K\cdot\left(D_T^K\right)^{-1}\cdot C_T^C\right]; \tag{28}$$

$$\widehat{\beta}_{T,1} = \left[A_T^C - B_T^C\cdot\left(D_T^C\right)^{-1}\cdot C_T^C\right]^{-1}\left[E_T^C - B_T^C\cdot\left(D_T^C\right)^{-1}\cdot C_T^K\right]. \tag{29}$$

Inverting equation (29) consequently gives:

$$\widehat{\beta}_{T,1}^{-1} = \left[E_T^C - B_T^C\cdot\left(D_T^C\right)^{-1}\cdot C_T^K\right]^{-1}\left[A_T^C - B_T^C\cdot\left(D_T^C\right)^{-1}\cdot C_T^C\right] \tag{30}$$

Simple transformations, $D_T^C = D_T^K \equiv D_T$, $E_T^C = E_T^K \equiv E_T$, $\widetilde{C}_T^K = D_T^{-1}C_T^K$, $\widetilde{C}_T^C = D_T^{-1}C_T^C$ then allow to rewrite (28) and (30):

$$\widehat{\gamma}_{T,1} = \left[\sum_{i=1}^{T} F_{T,i}^{K}{}^{-1}U_{T-i} - \sum_{i=1}^{T} F_{T,i}^{K}{}^{-1}X'_{T-i}\cdot\widetilde{C}_T^K\right]^{-1}\left[1 - \sum_{i=1}^{T} F_{T,i}^{K}{}^{-1}X'_{T-i}\cdot\widetilde{C}_T^C\right];$$

$$\widehat{\beta}_{T,1}^{-1} = \left[\sum_{i=1}^{T} F_{T,i}^{C}{}^{-1}U_{T-i} - \sum_{i=1}^{T} F_{T,i}^{C}{}^{-1}X'_{T-i}\cdot\widetilde{C}_T^K\right]^{-1}\left[1 - \sum_{i=1}^{T} F_{T,i}^{C}{}^{-1}X'_{T-i}\cdot\widetilde{C}_T^C\right].$$

Noting that $F_{T,i}^K$ and $F_{T,i}^C$ are defined in the following way proofs the proposition:

$$F_{T,i}^K \equiv (1-g)^{i-1}\cdot\sum_{s=1}^{T}(1-g)^{s-1}\cdot\frac{U_{T-s}}{U_{T-i}}y_{T-s}; \qquad F_{T,i}^C \equiv (1-g)^{i-1}\cdot\sum_{s=1}^{T}(1-g)^{s-1}\cdot\frac{y_{T-s}}{y_{T-i}}y_{T-s}.$$

■

Voting Mechanism. (Preliminary) Proof that the voting mechanism as given by equation (9) enforces revelation of the true desired policy.

To sketch the proof, define $I_1 \equiv [x_1, x^L), I_2 \equiv [x^L, x^H]$ and $I_2 \equiv (x^H, x_N]$ in an ordered sequence of desired policies as given by (8). Here x^L and x^H are defined assuming the agent told the truth. Moreover, define y as last periods inflation and x^* as the desired policy of a committee member. Consider the following three cases:

1. $x^* \in [x_1, x^L)$: The agent is able to influence the decision upwards by reporting a preferred policy outside the interval. However, this would move the policy contrary to his desired outcome.

2. $x^* \in [x^L, x^H]$: The agent is able to influence the decision only if he falsely claims to be prefer a policy that would constitute a new bound of the inactivity interval. However, this would ask him to over or to understate the current bound which then - if anything - would cause the decision outcome to be futher away from his preferred action. Hence, the the agent is again indifferent between reporting values that do not influence the decision but prefers truth telling over falsely repoting values that would have an impact.

3. $x^* \in (x^H, x_N]$: Here again, the only potential influence the agent can excert is to lower the upper bound of the inactivity interval. This however moves the policy away from what he prefers.

The argument assumes that agents know and take as given the preferred policies of the other committee members which is consistent with the informational assumptions of our model. Bargaining and explicit coordination is not considered. ■

B Model Uncertainty.

The perceived laws of motion are given by the following equation: $\widetilde{S}_t = A^n \cdot \widetilde{S}_{t-1} + B^n \cdot x_{t|t-1} + C^n \cdot u^n_{t1}$; Where $A^n = (A^n_0)^{-1} A^n_1, B^n = (A^n_0)^{-1} B^n_0$ and $C^n = (A^n_0)^{-1}$, for $n \in \{K, C\}$.

$$\widetilde{S}_t = \begin{bmatrix} U_t \\ U_{t-1} \\ y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ 1 \end{bmatrix}; \; A^C_0 = \begin{bmatrix} 1 & 0 & -\beta_1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \; A^K_0 = \begin{bmatrix} 1 & 0 & -\frac{1}{\gamma_1} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix};$$

$$A^C_1 = \begin{bmatrix} \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \; A^K_1 = \begin{bmatrix} -\frac{\gamma_2}{\gamma_1} & -\frac{\gamma_3}{\gamma_1} & -\frac{\gamma_4}{\gamma_1} & -\frac{\gamma_5}{\gamma_1} & -\frac{\gamma_6}{\gamma_1} & -\frac{\gamma_7}{\gamma_1} & -\frac{\gamma_0}{\gamma_1} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix};$$

$$B_0^{C,K} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} ; u_{t+1}^C = \begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ \xi_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} ; u_{t+1}^K = \begin{bmatrix} -\frac{1}{\gamma_1}\eta_{t+1} \\ 0 \\ \xi_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

A Bayesian decision-maker facing model uncertainty then constructs a simultaneous law of motion for the two models:

$$\begin{bmatrix} \widetilde{S}_{t+1} \\ \widetilde{S}_{t+1} \end{bmatrix} = \begin{bmatrix} A^C & 0 \\ 0 & A^K \end{bmatrix} \cdot \begin{bmatrix} \widetilde{S}_t \\ \widetilde{S}_t \end{bmatrix} + \begin{bmatrix} B^C \\ B^K \end{bmatrix} \cdot x_{t+1|t} + \begin{bmatrix} C_1^C & C_2^C \\ C_1^K & C_2^K \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_{t+1} \\ \xi_{t+1} \end{bmatrix}$$

Similarly the objective function caluculates losses predicted by both models simultaneously and attaches the committee member specific weight to each loss: $\rho = \begin{bmatrix} \widetilde{S}_{t+1} \\ \widetilde{S}_{t+1} \end{bmatrix}' Q_i \begin{bmatrix} \widetilde{S}_{t+1} \\ \widetilde{S}_{t+1} \end{bmatrix}$, where:

$$Q_i = \begin{bmatrix} w_t^i \cdot (M^{n\prime} \cdot Q^n \cdot M^n) & 0 \\ 0 & \left(1 - w_t^i\right) \cdot (M^{n\prime} \cdot Q^n \cdot M^n) \end{bmatrix}$$

and where M^n picks up the appropriate variables from the state vector and Q^n attaches the appropriate weights. Clearly $Q^C = Q^K$ and $M^C = M^K$ in the case of a single welfare criterion.

C Figures.

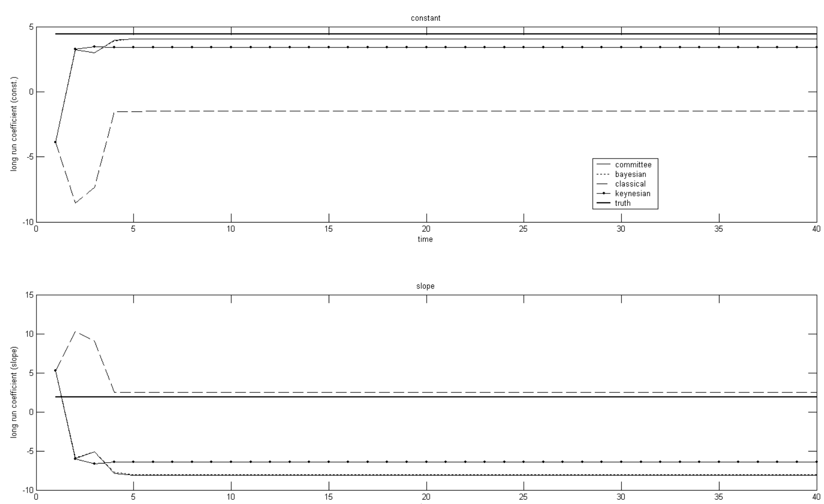

Figure (A) - Long Run Coefficients: OLS Estimates

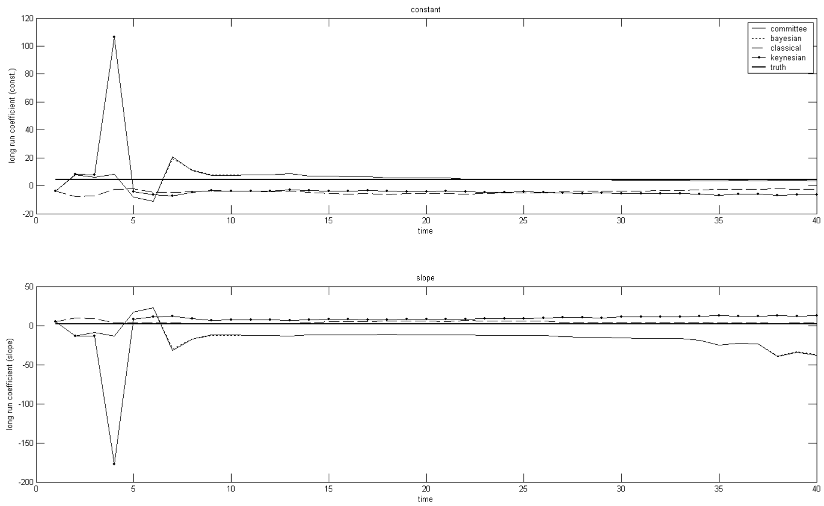

Figure (B) - Long Run Coefficients: Constant Gain Estimates

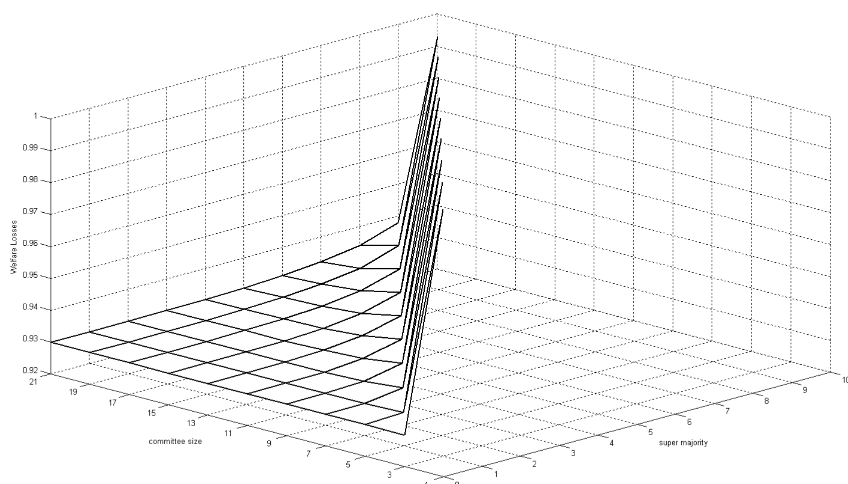

Figure (C) - Welfare Losses: Model A

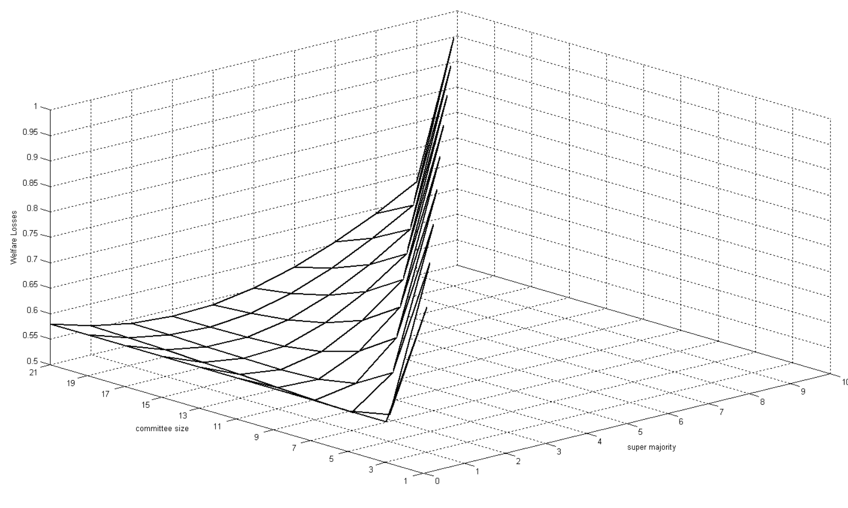

Figure (D) - Welfare Losses: Model B

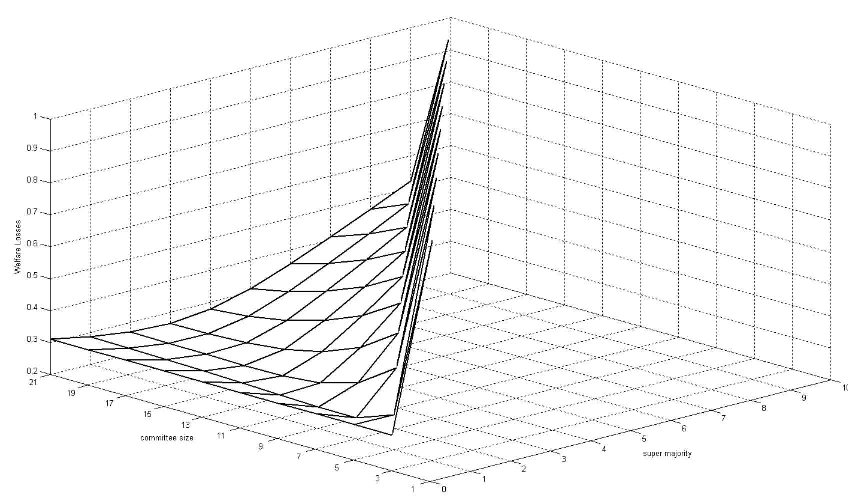

Figure (E) - Welfare Losses: Model C

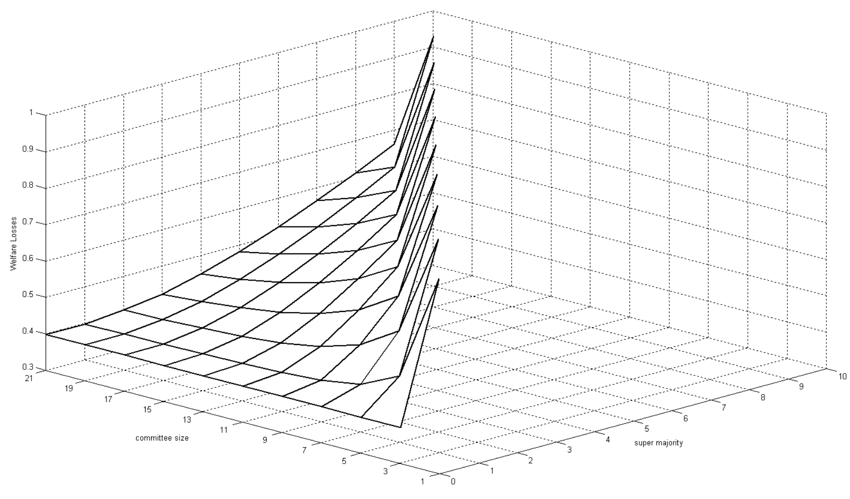

Figure (F) - Welfare Losses: Structural Break from Model B to Model A

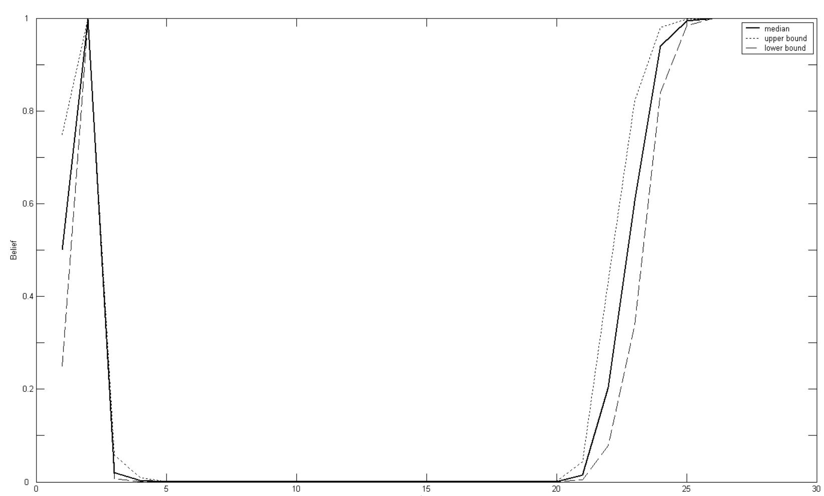

Figure (G) - Evolution of Median Beliefs: Model C

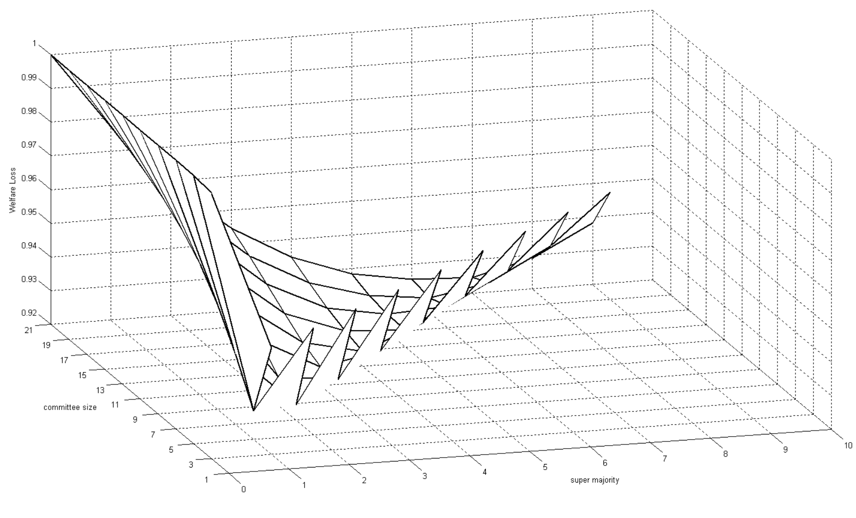

Figure (H) - Welfare Losses: Model B; Reduced Dispersion of Beliefs

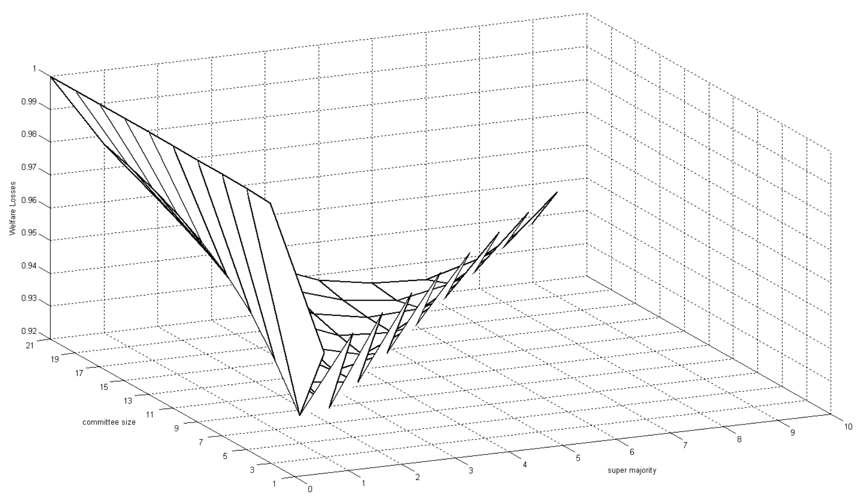

Figure (I) - Welfare Losses: Model B; "Crisis" (Volatile Economic Environment)