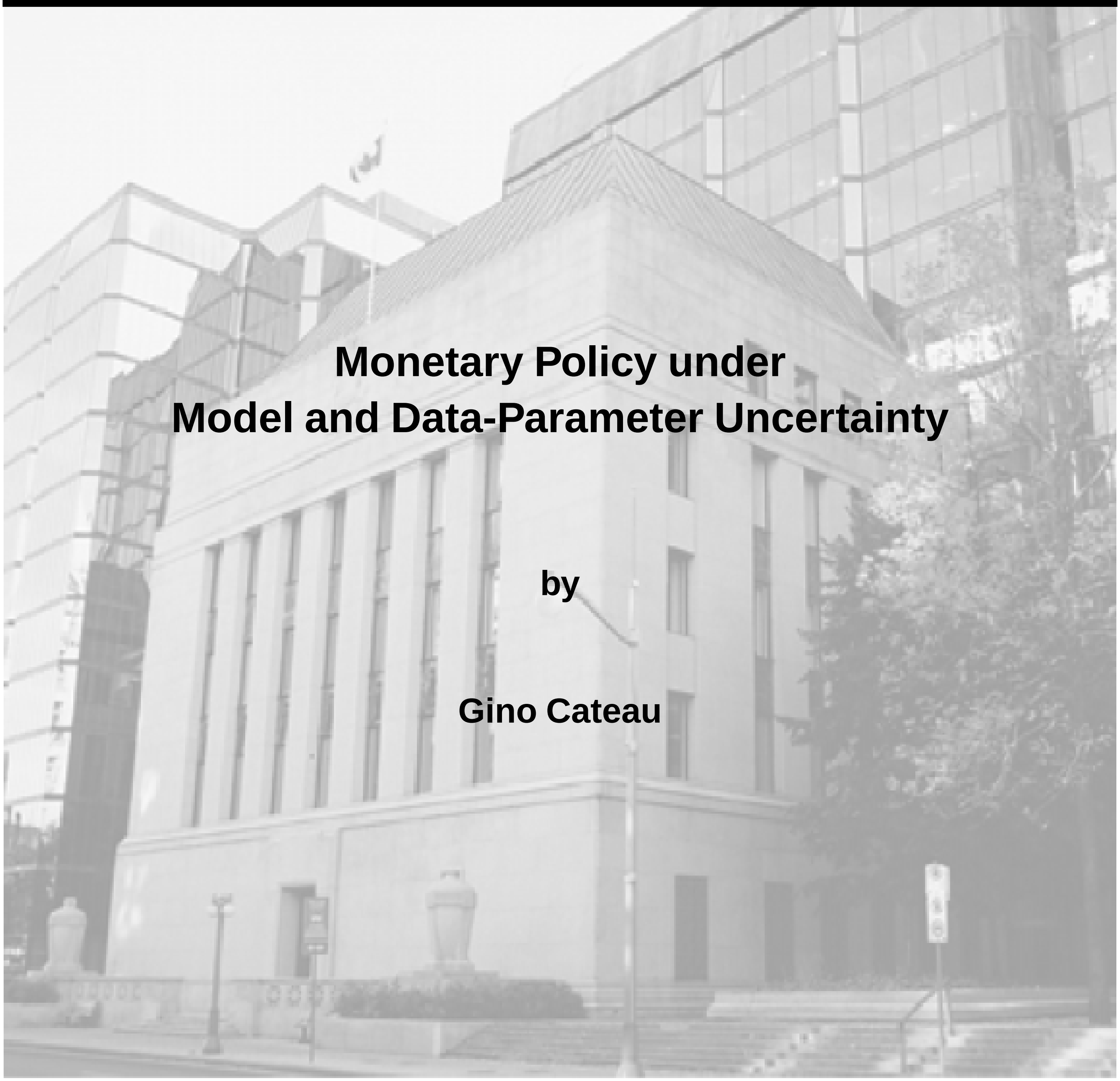
Bank of Canada Banque du Canada

Working Paper 2005-6 / Document de travail 2005-6



Monetary Policy under Model and Data-Parameter Uncertainty

by

Gino Cateau

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2005-6

March 2005

Monetary Policy under Model and Data-Parameter Uncertainty

by

Gino Cateau

Research Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
gcateau@bankofcanada.ca

Contents

Acknowledgements

I thank my advisers, Lars Peter Hansen, Peter Klibanoff, and Hanno Lustig, for their advice and comments. I thank Yuki Mano, Sonia Oreffice, Maral Kichian, Oleksiy Kryvtsov, Wendy Chan, and Manda Winlaw for their help and comments. I also thank seminar participants at the Bank of Canada, Federal Reserve Board, HEC Montréal, and Université Libres de Bruxelles. Remaining errors are mine.

Abstract

Policy-makers in the United States over the past 15 to 20 years seem to have been cautious in setting policy: empirical estimates of monetary policy rules such as Taylor's (1993) rule are much less aggressive than those derived from optimizing models. The author analyzes the effect of an aversion to model and data-parameter uncertainty on monetary policy. Model uncertainty arises because a central bank finds three competing models of the economy to be plausible. Data uncertainty arises because real-time data are noisy estimates of the true data. The central bank explicitly models the measurement-error processes for both inflation and the output gap, and it acknowledges that it may not know the parameters of those processes precisely (which leads to data-parameter uncertainty). The central bank chooses policy according to a Taylor rule in a framework that allows an aversion to the distinct risk associated with multiple models and data-parameter configurations. The author finds that, if the central bank cares strongly enough about stabilizing the output gap, this aversion generates significant declines in the coefficients of the Taylor rule, even if the bank's loss function assigns little weight to reducing interest rate variability. He also finds that an aversion to model and data-parameter uncertainty can yield an optimal Taylor rule that matches the empirical Taylor rule. Under some conditions, a small degree of aversion is enough to match the historical rule.

JEL classification: E5, E58, D8, D81
Bank classification: Uncertainty and monetary policy

Résumé

Depuis 15 à 20 ans, les autorités aux États-Unis semblent très prudentes dans la mise en œuvre de la politique monétaire : les valeurs estimées des paramètres des règles de politique monétaire telles que la règle de Taylor (1993) sont beaucoup moins élevées que celles issues de modèles d'optimisation. L'auteur analyse l'effet sur la politique monétaire du degré d'aversion de la banque centrale pour deux types d'incertitude : l'incertitude entourant la formulation appropriée du modèle et celle relative aux données et aux paramètres. Le premier type d'incertitude tient à l'existence de trois modèles plausibles de l'économie, et le second au fait que les données en temps réel sont des estimations imprécises des données véritables. La banque centrale modélise explicitement les processus d'erreur de mesure de l'inflation et de l'écart de production, mais elle est consciente qu'elle ne connaît pas leurs paramètres précis (d'où l'incertitude relative aux données et aux paramètres). Elle choisit sa règle de Taylor en tenant compte du risque spécifique associé à la multiplicité des configurations possibles en matière de modèles et de paramètres. L'auteur constate que, si la banque centrale est suffisamment préoccupée par la stabilisation de l'écart de production, il en résulte une diminution sensible des coefficients de la règle de Taylor même si la banque assigne une pondération peu élevée à la réduction de la variabilité des taux d'intérêt dans sa fonction de perte. Il relève également que l'aversion pour l'incertitude liée au choix du modèle et pour celle relative aux données et aux paramètres peut générer une règle de Taylor optimale conforme à la règle de Taylor empirique. Sous certaines conditions, un petit degré d'aversion suffit pour obtenir la règle empirique.

Classification JEL : E5, E58, D8, D81
Classification de la Banque : Incertitude et politique monétaire

1. Introduction

Since the work of Taylor (1993), the monetary policy literature has focused increasingly on characterizing desirable monetary policy in terms of simple interest rate feedback rules; i.e., guidelines where the central bank adjusts the interest rate, its policy instrument, in response to economic conditions. Typically, given some structural model that links inflation, output, and the interest rate, the central bank sets policy according to the interest rate feedback rule to minimize a weighted average of inflation, output, and interest rate variability. The result of that exercise is often puzzling, however: unless the central bank assigns an implausibly high weight to smoothing the interest rate,[1] parameters of optimal policy rules call for much stronger responses to inflation and output than those estimated from historical data.[2]

What can explain this apparent reluctance of policy-makers to act aggressively? One important branch of this literature argues that attenuated policy is the result of policy-makers facing uncertainty, whether regarding the model parameters or the data. It has generally been accepted, since the work of Brainard (1967), that parameter uncertainty can lead to less aggressive policy. Brainard's intuition indeed seems sensible: strong responses required by optimal policy in the context of a particular model are dependent upon the parameters of that model being known. Because these parameters are in fact not known, but need to be estimated, taking account of one's uncertainty about them justifies gentler responses. However, most studies that formally incorporate parameter uncertainty in the central bank's decision process find that it has a negligible effect on policy; i.e., it does not lead to a significant attenuation[3] of the policy rule parameters (Estrella and Mishkin 1999, Peersman and Smets 1999).

Similarly, although considerable differences can exist between real-time and final estimates of inflation and the output gap (Orphanides 2001), various authors have found that sensible degrees of measurement error do not lead to a high enough attenuation in the policy rule parameters (Rudebusch 2001).

The papers cited above assume no model uncertainty: although the central bank is unsure about the model parameters or fears data uncertainty, it is confident that the structure of the model is the right one for policy-making. A related strand of the literature examines whether a direct concern for model uncertainty can help to explain why policy-makers may prefer less aggressive policy. Much of that literature assumes that policy-makers have one good reference model for setting policy but are concerned about uncertain possible deviations from it. Therefore, they use a robust control approach (Hansen and Sargent 2004, Onatski and Stock 2002) to design policy rules that resist deviations from their particular reference model. But what if the central bank is uncertain about competing *reference* or *non-nested* models of the economy?

This paper considers a central bank that faces competing reference models of the economy.[4] In the first part of the paper, the central bank's problem is to choose a Taylor rule which works reasonably well given its difficulty in deciding between the various models that it finds plausible.

There are several ways in which such a rule can be chosen. The central bank, for instance, can take a worst-case model approach. A drawback of this approach is that it can make the central bank very conservative. Moreover, it disregards any information or beliefs that the central bank may have regarding the plausibility of a particular model. Alternatively, if the central bank can assign weights that reflect the plausibility of each model, it can choose a rule that minimizes the average of the different models' losses. A benefit of this approach is that a rule can be allowed to perform poorly in a particular

[1] Rudebusch (2001) finds, in the context of his model, that if the central bank assigns *five* times more weight to controlling interest rate variability than it assigns to controlling inflation and output-gap variability, the optimal Taylor rule can match the historical one.

[2] This result is found, for instance, in Rudebusch (2001), and Tetlow and von zur Muehlen (2001).

[3] Some studies that use an unrestricted vector autoregression to describe the economy and an unrestricted policy rule report more important quantitative effects of parameter uncertainty (Soderstrom 1999, Sack 2000).

[4] Levin and Williams (2003) seek to identify a robust policy rule when policy-makers face competing reference models. My motivation is different. I try to determine whether uncertainty across models can explain why policy-makers may prefer less aggressive policy rules. Moreover, I differ in my approach to handling model uncertainty.

model as long as that model is not very plausible. A disadvantage is that the central bank chooses a rule as if average performance is the only criterion that matters; there is no notion that the central bank may, per se, care about preventing bad performances in any particular model (i.e., it may want to achieve some robustness as well).

This paper uses a framework that exhibits the non-reduction of two-stage lotteries (Segal 1990, Klibanoff, Marinacci, and Mukerji 2002, Ergin and Gul 2004) to account for model uncertainty. What this means, in my context, is that the central bank distinguishes between two distinct kinds of risks: a first-order risk, which arises given a particular model and its stochastic properties, and a second-order risk, which is associated with the multiplicity of models. Similar to Klibanoff, Marinacci, and Mukerji (2002), I interpret the attitude of the central bank towards that second-order risk as the attitude towards model uncertainty. Just as decision makers can be risk-averse, neutral, or loving in classical risk theory, my framework allows the central bank to be model-uncertainty averse, neutral, or loving. And as in Klibanoff, Marinacci, and Mukerji (2002), my framework nests both the worst-case and Bayesian approaches as special cases: the worst-case approach is the limiting case, where the degree of aversion to model uncertainty tends to infinity (see section A.1 of Appendix A), while the average loss approach obtains when the central bank is neutral to model uncertainty.

Specifically, the policy problem I analyze in this paper considers a central bank that views the models of Fuhrer and Moore (1995), Rudebusch and Svensson (1999), and Woodford (1996, 1999) (extended by Giannoni 2000) as three plausible models of the economy. Woodford and Giannoni's model is a theoretically founded forward-looking model; i.e., it is derived by log-linearizing the Euler equations of firms and consumers in the economy. Rudebusch and Svensson's model is an ad hoc backward-looking model, but it is designed to be a good empirical model. Fuhrer and Moore's model, on the other hand, incorporates both forward- and backward-looking elements and provides a good description of empirical features such as inflation persistence. My objective is to analyze how the rule chosen changes when I vary the central bank's degree of aversion to model uncertainty. I find that policy becomes less aggressive as I increase the degree of aversion to model uncertainty. Moreover, if the central bank cares strongly enough about stabilizing the output gap, this aversion generates important declines in the coefficients of the Taylor rule *even* when the central bank's loss function gives little weight to reducing interest rate variability.

In the second part of the paper, I extend the policy problem to one where the central bank is not only uncertain about the three competing models, but also considers data uncertainty to be important. The central bank accounts for data uncertainty by modelling the measurement-error processes for the output gap and inflation, but recognizes that it is uncertain about the parameters of those processes (henceforth, this is referred to as data-parameter uncertainty). The decision-making framework is extended to incorporate a second-order risk when evaluating policy across models *and* parameter configurations. I find that, in the presence of data-parameter uncertainty as defined above, an increase in aversion to model and data-parameter uncertainty can generate an optimal Taylor rule that matches the empirically observed Taylor rule.

I interpret the economic significance of the degree of aversion to model uncertainty by relating it to the proportional premium that the central bank would pay to be indifferent between facing model uncertainty or achieving the average loss of the models for sure. When the central bank assigns a higher weight to output stabilization than to reducing interest rate variability, I find that small degrees of aversion are enough to generate relatively important declines in the responses to inflation and the output gap, provided that the central bank believes more in Woodford and Giannoni's model. I then show that, for the benchmark case where the central bank assigns a higher weight to the Woodford and Giannoni model than to the other two models, a small degree of aversion is enough to generate an optimal Taylor rule that matches the empirical Taylor rule.

The rest of this paper is organized as follows. Section 2 describes a general framework for handling model and parameter uncertainty. Section 3 analyzes the monetary policy problem under model uncertainty. Section 4 expands the policy problem to analyze a concern about model-parameter, data uncertainty, and data-parameter uncertainty. Section 5 provides an interpretation of the economic significance of the model uncertainty aversion parameter. Section 6 relates my result to the literature and section 7 concludes. Derivations, the analysis for a generalized Taylor rule, figures and tables are provided in the appendixes.

2. A Decision-Making Framework

2.1 Decision making under model uncertainty

Suppose that the true data-generating model is G, but that the decision maker does not know it. Not only does the decision maker not know the true model, but suppose that, maybe because they are faced with competing theories suggesting different models, they find it difficult to settle on a particular model. As a result, suppose that the decision maker considers $\{G_k, k = 1, ..., n\}$ as the set of plausible models. Although the agent finds model $k = 1, ..., n$ plausible, their degree of belief in each may vary. Assume that π_k is the weight that the agent attaches to the relevance of model k such that $\sum_k \pi_k = 1$.

Moreover, suppose that the agent is also given a set of feasible rules,

$$\{K(\gamma), \gamma \in \Gamma\}, \tag{1}$$

and a loss function,

$$v_k(\gamma) = V(G_k, K(\gamma)),$$

which measures the loss of applying the rule $K(\gamma)$ when the model is G_k. The agent's objective is to find *one* rule, say $K(\hat{\gamma})$, that minimizes their loss given their inability to decide between the competing models.

One possible method of designing a rule that works reasonably well is to minimize the following objective function with respect to γ:

$$wc(\gamma) = \max \{v_1(\gamma), v_2(\gamma), ..., v_n(\gamma)\}. \tag{2}$$

Thus, $wc(\gamma)$ evaluates the loss of the rule $K(\gamma)$ under the worst-case model. As a result, the rule is expected to work reasonably well no matter which model in the set turns out to be relevant. This approach has two main problems. First, it does not, in general, separate the effect of a change in attitude towards model uncertainty from the effect of a change in model uncertainty in the environment. To see this, compare the rule chosen when the agent considers n models to the rule chosen when they instead consider $n' > n$ models. If the rules differ, unless the agent's attitude is fixed to the extreme one where they care only about the worst-case scenario, it cannot be said for certain whether the change in their behaviour is due to the fact that they now contemplate more models or whether their attitude to having more models is now different. Thus, in the worst-case approach, I can single out the effect of an aversion to model uncertainty only if I fix the agent's attitude to the most extreme one. But why should an agent be so conservative?

The second drawback is that the worst-case approach disregards any information or beliefs that the agent may have regarding the plausibility of a particular model. With a criterion like (2), the agent can be led to choose a corner solution that is quite restrictive even though they may give very little weight to the models that lead to the corner solution.

An alternative approach, which incorporates the beliefs that the agent may have relative to the relevance of a particular model, is to choose a rule that minimizes the average loss of the models; i.e., a Bayesian approach. Given the plausibility weights attached to each model, this approach implies that the agent chooses γ to minimize

$$av(\gamma) = \sum_{k=1}^{n} \pi_k v_k(\gamma). \tag{3}$$

This approach has the advantage that the least plausible models are given the least weight in the decision process. Therefore, it is consistent with the idea that a rule that does not perform very well in a particular model is permissible if that particular model is not very plausible; what matters most is that the rule performs reasonably well in the models that are more likely to be relevant. A disadvantage, however, is that there is no notion that the agent may care about something more than the average performance. In particular, there is no notion that the agent may be unwilling to accept rules that yield very bad performances in some models even though they may perform well on average. Therefore, there is no notion that the agent may, per se, want to achieve some robustness.

To avoid the drawbacks of the above two alternatives, I assume that the agent cares about models in a setting that derives from the literature on decision theory. Specifically, I assume that the agent chooses γ to minimize

$$h(\gamma) = \sum_{k=1}^{n} \pi_k \phi(\ v_k(\gamma)\ ). \tag{4}$$

The objective in (4) has two noticeable features. First, as in the Bayesian approach, the plausibility of each model matters in the decision process. Second, contrary to the Bayesian approach, the objective function is an average of each model's loss *transformed* by ϕ. What role does the function ϕ play? Segal (1990), Klibanoff, Marinacci, and Mukerji (2002), and Ergin and Gul (2004) develop a decision theory that analyzes preference orderings for which compound lotteries cannot be reduced to simple lotteries. In terms of the problem at hand, I can think of (4) in similar terms. The decision maker faces a first-order risk when they evaluate the rule γ in a particular model given the stochastic properties of the model,[5] but they also face a second-order risk when they evaluate γ across models. Since ϕ is in general non-linear, this means that, in general, the compound lottery cannot be reduced to a simple lottery. As in Klibanoff, Marinacci, and Mukerji (2002), ϕ further captures the attitude of the agent towards model uncertainty. In particular, since I seek to minimize (4), a convex ϕ implies aversion to model uncertainty, a concave ϕ implies model-uncertainty love, and a linear ϕ implies model-uncertainty neutrality. When ϕ is linear, (4) collapses to (3). So the Bayesian approach is the special case where the agent is neutral to model uncertainty (the compound lottery can be reduced to a simple lottery). On the other hand, as I show in section A.1 of Appendix A, the worst-case approach (2) is the limiting case where the degree of aversion to model uncertainty tends to infinity. Conversely, if the degree of model-uncertainty love goes to infinity, then the agent makes decisions according to a best-case scenario.

2.2 Decision making under model and parameter uncertainty

Suppose that not only is the agent uncertain across models, they are also uncertain about some parameters in each model.[6] Therefore, suppose that model k depends on some parameters, $\theta_k \in \Theta_k$, that are not known to the agent (i.e., $G_k(\theta_k)$) and suppose that the agent puts a prior $\mathcal{P}_k$ on Θ_k to capture their uncertainty about θ_k.

Define

$$v_k(\theta_k, \gamma) = V(G_k(\theta_k), K(\gamma))$$

as the indirect loss of applying the rule $K(\gamma)$ when the model is G_k. In view of this model and parameter uncertainty, how should the agent optimally choose γ? To handle both model and parameter uncertainty,

[5] Although I do not make it implicit, it is understood in this section that v_k is stochastic. Indeed, v_k is a weighted average of unconditional variances determined from model k. Hence it is stochastic.

[6] In the applications that follow, parameter uncertainty arises because the central bank is uncertain about the parameters of the measurement-error processes.

I extend (4) to a framework where the agent chooses γ according to:

$$h^*(\gamma) = \sum_{k=1}^{n} \pi_k \, E_{\mathcal{P}_k} \left(\, \phi(\, v_k(\theta_k, \gamma) \,) \, \right). \tag{5}$$

In this setting, the agent faces a second-order risk when they evaluate γ across various models and parameter configurations. ϕ, therefore, captures the attitude of the agent towards *both* model *and* parameter uncertainty. Indeed, as I show in section A.2 of Appendix A, if I increase the degree of aversion (defined as $\frac{\phi''}{\phi'}$) towards infinity, (5) collapses to the worst-case criterion:

$$wc^*(\gamma) = \max \left\{ \max_{\theta_1 \in \Theta_1} v_1(\theta_1, \gamma), \max_{\theta_2 \in \Theta_2} v_2(\theta_2, \gamma), ..., \max_{\theta_n \in \Theta_n} v_n(\theta_n, \gamma) \right\}.$$

Therefore, ϕ captures how the agent reacts to having many models and parameter configurations to consider for decision making. For a high enough curvature in ϕ, the agent protects themself against their uncertainty about models and parameters by making decisions according to the worst-case model and parameter configuration. Evidently, if ϕ is linear, (5) collapses to a Bayesian criterion, where the average is taken with respect to the priors over models and parameters.

3. Monetary Policy under Model Uncertainty

Taylor observed in 1993 that, since Alan Greenspan became chairman of the U.S. Federal Reserve in 1987, monetary policy in the United States could be very well described by a simple policy rule where the federal funds rate was increased by 1.5 per cent in response to a 1 per cent change in average inflation (over the past four quarters) and by 0.5 per cent in response to a 1 per cent change in the output gap (the deviation of actual real GDP from potential GDP); i.e.,

$$i_t = 1.5\bar{y}_t + 0.5x_t, \tag{6}$$

where i_t denotes the (de-meaned) federal funds rate, $\bar{y}_t = \sum_{j=0}^{3} y_{t-j}$ is the (de-meaned) inflation rate over the past four quarters, and x_t is the output gap. Various authors have since re-estimated Taylor's rule over longer time periods and using more sophisticated econometric techniques. Their results are broadly consistent with Taylor's (see Rudebusch 2001 and the references therein). As a crude benchmark, historical estimates of U.S. monetary policy during the late 1980s and 1990s suggest that policy can be reasonably well described by a Taylor rule with an inflation response between 1.5 and 2 per cent and an output-gap response between 0.5 and 1 per cent.

Given the empirical robustness of the Taylor rule, various authors have tried to answer the following question: Can an optimal Taylor rule match its empirical counterpart? Typically, the authors assume that the central bank has *one* well-defined model and that its problem is to optimally choose the coefficients of a Taylor rule to minimize some weighted average of inflation, output-gap, and interest rate variability. The results of these investigations are often puzzling, however: optimal Taylor rules often require much more aggressive policy than their empirical counterparts; i.e., optimal Taylor rules yield coefficients that are often much bigger than their empirical estimates.

To make sense of this puzzle, three main arguments have been offered. First, some authors consider Brainard-style stories, where the central bank is uncertain about the parameters of its model. Second, some authors argue that the central bank may fear data uncertainty. Third, others argue that the central bank cares strongly about limiting interest rate variability; i.e., the central bank assigns a high weight to the variability of the interest rate in the objective function that it minimizes. I believe that these reasons are not totally satisfactory. On the one hand, various authors find that parameter uncertainty and plausible degrees of data uncertainty have a very small effect on optimal policy (Rudebusch 2001, Estrella and Mishkin 1999), whereas, on the other hand, it is difficult to theoretically rationalize why the central bank should place such a high weight in controlling the interest rate's variability (Woodford 1999, Svensson 2003).

In this section, I investigate whether uncertainty across non-nested models can justify why the central bank may prefer less aggressive policy. In particular, I consider a central bank that sets policy according to a Taylor rule,[7]

$$i_t = \psi_y \bar{y}_t + \psi_x x_t, \tag{7}$$

and which is undecided between three models of the economy: Woodford (1999) and Giannoni (2000) (henceforth, the WG model), Rudebusch and Svensson (1999) (henceforth, the RS model), and Fuhrer and Moore (1995) (henceforth, the FM model).

The WG model is a forward-looking model derived from explicit microfoundations. It consists of an intertemporal IS equation that relates the output gap to the interest rate, expected inflation, and the natural rate of interest, r_t^n; i.e., the interest rate that would prevail in the absence of any distortions,

$$x_t = E_t x_{t+1} - \sigma^{-1}(i_t - E_t y_{t+1} - r_t^n), \tag{8}$$

and the aggregate supply equation (or expectational Phillips curve), which relates inflation to the output gap, expected inflation, and a supply shock, u_t:

$$y_t = \kappa x_t + \beta E_t y_{t+1} + u_t. \tag{9}$$

Note that output and inflation are purely forward looking; that is, they do not display any intrinsic persistence. Therefore, any persistent movement in output and inflation must come from the exogenous shocks hitting the economy. For convenience, the WG model assumes that r_t^n and u_t^n are AR(1) processes:

$$r_{t+1}^n = \rho_r r_t^n + \epsilon_{r,t+1}, \tag{10}$$

$$u_{t+1} = \rho_u u_t + \epsilon_{u,t+1}, \tag{11}$$

where $\epsilon_{r,t+1}$ and $\epsilon_{u,t+1}$ have mean 0 and variances σ_r^2, σ_u^2, respectively.

The RS model, on the other hand, is purely backward looking. It is not derived from first principles, but it is designed to be a good empirical model, in that it has rich dynamics and is claimed to be similar in structure to models that central bankers actually use. The model consists of an aggregate supply equation that relates inflation to the output gap and lags of inflation,

$$y_{t+1} = b_0 y_t + b_1 y_{t-1} + b_2 y_{t-2} + (1 - b_0 - b_1 - b_2) y_{t-3} + b_x x_t + e_{y,t+1}, \tag{12}$$

and an aggregate demand equation that relates the output gap to the interest rate,

$$x_{t+1} = d_0 x_t + d_1 x_{t-1} - d_i(\bar{i}_t - \bar{y}_t) + e_{x,t+1}, \tag{13}$$

where $\bar{y}_t = \frac{1}{4}\sum_{j=0}^{3} y_{t-j}$ is average quarterly inflation, $\bar{i}_t = \frac{1}{4}\sum_{j=0}^{3} i_{t-j}$ is the average quarterly federal funds rate, and $e_{y,t+1}$, $e_{x,t+1}$ have mean 0 and variances $\sigma_{e_y}^2$, $\sigma_{e_x}^2$, respectively. Therefore, in contrast to the WG model, both output and inflation display intrinsic persistence in a completely backward-looking specification.

The FM model, on the other hand, combines both forward- and backward-looking elements. FM assume an economy where agents negotiate nominal-wage contracts that remain in effect for four quarters. FM assume a fixed markup from wages to price such that there is little distinction between wages and price. Inflation is defined as:

$$y_t = 4(p_t - p_{t-1}), \tag{14}$$

[7] See Appendix B for an analysis with a Taylor rule that allows for interest rate inertia.

and the price index, p_t, is defined as a moving average of current and past nominal contract prices, w_t:

$$p_t = \sum_{i=0}^{3} f_i w_{t-i}, \tag{15}$$

where the weight, f_i, is given by $f_i = 0.25 + (1.5 - i)s$. The real contract price index, v_t, is defined as the weighted average of current and past real contract prices, $w_t - p_t$:

$$v_t = \sum_{i=0}^{3} f_i(w_{t-i} - p_{t-i}). \tag{16}$$

Agents determine the current real contract price as a function of the real contract prices that are expected to prevail over the duration of the contract, adjusted for excess demand conditions and an identically, independently distributed (i.i.d.) shock,

$$w_t - p_t = \sum_{i=0}^{3} f_i E_t(v_{t+i} + r\, x_{t+i}) + \epsilon_{w,t}. \tag{17}$$

The aggregate demand relation makes the output gap a function of its own lags and the *ex-ante long-term* real interest rate, ρ_t:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} - a_\rho \rho_{t-1} + \epsilon_{x,t}. \tag{18}$$

Finally, ρ_t is defined, according to the pure expectations' hypothesis, as a weighted average of the future short-term real interest rate,

$$\rho_t = d\, E_t \rho_{t+1} + (1-d)(i_t - E_t y_{t+1}). \tag{19}$$

Of course, these three models represent only a very small subset of the entire set of competing representations of the U.S. economy. Nevertheless, I restrict my analysis to these three models because they differ markedly in terms of their construction, assumptions, and predictions. Model uncertainty, I believe, is likely to be a thornier issue when the central bank has models at hand that are built from different assumptions and lead to different predictions, as is the case here.

3.1 A modified optimization problem

The monetary policy literature usually assumes that the objective of the central bank is to choose policy to minimize some weighted average of the unconditional variance of inflation, the output gap, and the change in the interest rate:

$$\text{var}(y_t) + \omega \text{var}(x_t) + \nu\ \text{var}(i_t - i_{t-1}). \tag{20}$$

Therefore, if the central bank had *one* model, its problem would be to optimally pick the coefficients of (7) to minimize (20). Put differently, given the model, the central bank would compute the unconditional variances of inflation, output gap, and interest rate for each (ψ_y, ψ_x), and, given the weights ω and ν, it would pick the (ψ_y, ψ_x) combination that yields the minimum weighted average of the unconditional variances.

The problem at hand differs from the above in that the central bank has three non-nested models of the economy, and hence three distinct ways of computing the unconditional variances of inflation, output, and the interest rate. As motivated in section 2, I assume that the central bank handles model uncertainty in the framework provided by (4).

Denote the coeffients of the Taylor rule as $\gamma = (\psi_y, \psi_x)$, and denote $v_{WG}(\gamma)$ as the indirect loss of applying (7) to minimize (20) when the unconditional variances are computed according to the WG model. Define $v_{RS}(\gamma)$ and $v_{FM}(\gamma)$ likewise. Then, denoting π_{WG}, π_{RS}, and π_{FM} as the probabilities that the central bank assigns to the plausibility of the WG, RS, and FM models, respectively ($\pi_{WG}+\pi_{RS}+\pi_{FM}=1$), the central's bank problem is to choose γ to minimize

$$\pi_{WG}\,\phi(v_{WG}(\gamma))+\pi_{RS}\,\phi(v_{RS}(\gamma))+\pi_{FM}\,\phi(v_{FM}(\gamma)). \tag{21}$$

For the remainder of this paper, I will assume that ϕ is the increasing function

$$\phi(x)=\frac{e^{\eta x}}{\eta},\quad \eta\neq 0. \tag{22}$$

Equation (22), or any affine transformation of it, represents constant model-uncertainty aversion when $\eta>0$, or constant model-uncertainty love if $\eta<0$. Informally, a constant model-uncertainty attitude means that if the loss of each model is changed by the same constant amount, the behaviour of the decision maker is unaffected: when the decision maker displays constant model-uncertainty aversion, what matters is the relative expected losses, not the level of the losses. Hence, changing the losses of each model by the same amount does not change the implicit weight given to each model in the decision process. When $\eta=0$, (22) collapses[8] to $\phi(x)=x$. Therefore, $\eta=0$ corresponds to model-uncertainty neutrality.

My objective is to address the question: Does a higher degree of model-uncertainty aversion make monetary policy more or less aggressive? To do this, I investigate the impact of varying η, the degree of model-uncertainty aversion on the coefficients, γ, of the Taylor rule. Moreover, I investigate the effect of varying the plausibility weight that the central bank attaches to each model.

3.2 Solution and computation details

To solve the optimization problem, I first need to

(i) specify the parameters of the WG, RS, and FM models,

(ii) specify the weights assigned to the output-gap and interest rate stabilization in the loss function; i.e., ω,ν, and

(iii) solve for the indirect losses $v_{WG}(\gamma), v_{RS}(\gamma)$, and $v_{FM}(\gamma)$ for each γ.

I set the parameters of the WG, RS, and FM models according to Tables 1, 2, and 3, respectively. These values are consistent with the authors' original calibrations.

Determining the weights that the central bank attaches to output stabilization and interest rate variability control is a difficult exercise. Calibration exercises that attempt to pin down those values are done under the assumption that the calibrated model is the right one for the economy. This is not the case for the problem at hand, since the central bank in this study faces model uncertainty. It is not clear that a central bank that faces model uncertainty will assign the same weights to output stabilization and interest rate variability control as one that does not.

Even disregarding the above issue, however, there is considerable disagreement in the literature about what the weights to output-gap stabilization and interest rate variability control should be. For instance, when derived from models with microfoundations, the weight on output stabilization is quite sensitive to the type of staggered price-setting behaviour one assumes. Indeed, Woodford (2000) finds that, when sticky prices are adjusted at random intervals, as in Calvo (1983), the weight on output stabilization is quite small ($\omega\approx 0.01$). On the other hand, Erceg and Levin (2002) find a weight of $\omega\approx 1$ when prices are determined by staggered nominal contracts of fixed duration, as in Taylor (1980). Similarly, the weight that should be assigned to reducing the interest rate variability is not without controversy.

[8]Since ϕ is unique up to affine transformations, taking the limit of $\phi(x)=\frac{e^{\eta x}-1}{\eta}$ as η goes to 0.

Svensson (2003), for instance, argues that it is difficult to motivate theoretically why a central bank should have a concern for controlling the interest rate variability. Paradoxically, empirical studies find that to match U.S. data, a much higher weight must be assigned to reducing the interest rate variability than to output-gap stabilization. For example, Soderstrom, Soderlind, and Vredin (2002) find that their small New Keynesian model of monetary policy can match U.S. data if the weight assigned to output-gap stabilization is small ($\omega \leq 0.1$), while the weight assigned to interest rate variability control is much bigger ($0.5 \leq \nu \leq 2$). Similarly, Dennis (2003) finds that the weight assigned to output stabilization is not statistically different from zero. In view of those difficulties, I conduct my analysis under the grid of values given by $\omega = 0.1, 0.5, 1$ and $\nu = 0.1, 0.5, 1$.

Computing the indirect loss of the RS model is not very difficult. Since the model is purely backward looking, once it is put in state-space form, the unconditional variance of the state vector is the fixed point of a Sylvester equation. I use the doubling algorithm, as outlined in Hansen and Sargent (2004), to solve for this fixed point. The indirect loss can then be constructed from this fixed point, as I show in section C.1 of Appendix C. For the WG and FM models, the solutions are more complicated because they are forward looking, and what agents know and how they form expectations matters. I make the following assumptions regarding private agents' information:

Assumption 1 *Agents face no model uncertainty.*

Assumption 2 *For any t, agents know the central bank's Taylor rule, $i_t = \psi_y \bar{y}_t + \psi_x x_t$.*

I impose Assumption 1 to stress that I am primarily interested in the central bank's decision-making problem in the face of model uncertainty. Agents behave in a particular way in the economy and the very reason the central bank faces model uncertainty is that it is unsure about how to model that behaviour. As a result, it postulates three alternative models of agents' behaviour: the WG, RS, and FM models. As in Sargent (1999), Assumption 2 is useful because it allows me to eliminate state variables that would otherwise come from the agents' optimization problem. Furthermore, it can be rationalized as the outcome of a system where central bank watchers give agents an accurate estimate of the central bank's decision rule. Alternatively, I can assume that the central bank announces its decision rule to the public and commits to it. Though strong, this assumption retains the flavour that the central bank has to assess agents' expectations to set policy.

To solve for $v_{WG}(\gamma)$, I first need to compute agents' expectations of inflation and output. I show in section C.2.1 of Appendix C how to derive these. Given the expectations of inflation and output and the Taylor rule (7), I can find the transition equation for the predetermined state variables. The unconditional variance of the predetermined state variables is the fixed point of a Sylvester equation. $v_{WG}(\gamma)$ is then constructed from this fixed point as shown in section C.2.1 of Appendix C. Finally, to solve for $v_{FM}(\gamma)$, I use the Anderson-Moore algorithm (Anderson and Moore 1985) which uses an invariant subspace method to solve the Fuhrer and Moore model.

Finally, I solve for the optimal γ by using the Nelder-Mead direct search algorithm.

3.3 Results

3.3.1 Optimal vs empirical Taylor rules: a puzzle

To give substance to the motivation of this paper, I first take each model in isolation and compute the optimal Taylor rule across various configurations of the central bank preferences towards output stabilization and interest rate variability control. Table 4 displays the optimal Taylor rules for the WG, RS, and FM models as I vary ω and ν in turn between 0.1, 0.5, and 1, respectively. I first notice that, across all the central bank preferences, the optimal Taylor rules yield responses to inflation, ψ_y, that are bigger than 2 for all three models. Thus, in response to a change in inflation, each of the three models requires changing the interest rate more agressively than what an empirically measured Taylor rule suggests. But among the three models, the WG model requires responses to inflation that are much more aggressive than those required by the RS and FM models; indeed, the WG model requires responses to inflation that are 2 to 3 times higher than those required by the RS or FM models.

On the other hand, the RS model requires optimal responses to the output gap that are bigger than 1 across all of the central bank's preferences, except when $\omega = 0.1$ and $\nu = 1$, where it yields a response of 0.93. The other two models are more sensitive to the weight that the central bank assigns to output-gap stabilization and the control of interest rate variability. Indeed, both the WG and FM models require aggressive responses to the output gap when the central bank cares more about stabilizing the output gap than controlling the variability of the interest rate; i.e., $\omega > \nu$. But the responses to the output gap can decline below 1 or even be close to zero, if the central bank assigns a higher weight to controlling the interest rate variability than it does to output-gap stabilization. However, even though the WG and FM models can generate responses to the output gap that are quite low for some configurations of ω and ν, these are coupled with responses to inflation that are too aggressive, especially for the WG model.

Therefore, Table 4 illustrates that, for the three models under consideration, the main reason optimal Taylor rules fail to match the empirical Taylor rule is that the models require overaggressive responses to changes in inflation. Table 4 also illustrates that a high weight to reducing interest rate variability seems to be necessary to yield optimal Taylor rules that come close to the empirically observed one. Indeed, I find that, when $\omega = 1$, increasing ν yields optimal Taylor rules that get progressively closer to the empirical one, at least for the FM and RS models. As Svensson (2003) argues, however, it is difficult to theoretically justify why the central bank should assign such a high weight to interest rate variability control. Finally, Table 4 illustrates that the three models considered require very different policy actions to minimize the central bank's loss function. The WG model requires very strong interest rate changes with respect to changes in inflation. In contrast to the WG model, the RS and FM models require milder interest rate changes with respect to inflation, but the RS model requires stronger interest rate changes than the FM model with respect to changes in the output gap. In view of the magnitudes involved, the attitude of the central bank towards model uncertainty is likely to have important effects on the choice of the optimal rule, since the sensitivity of each model to deviations from its optimum becomes very important. I analyze this sensitivity below, as well as the effect of an aversion to model uncertainty on the choice of the Taylor rule.

3.3.2 The optimal Taylor rule under model uncertainty

Suppose, at the outset, that the central bank gives more importance to stabilizing output than to controlling the interest rate's variability. Figure 1 plots the loss surface of the WG model; i.e., $v_{WG}(\gamma)$ over the (ψ_y, ψ_x) plane when the weights to output stabilization and interest rate variability control are, respectively, $\omega = 1$ and $\nu = 0.5$. The loss surface reaches a minimum of 9.77 at $(\psi_y^{WG}, \psi_x^{WG}) = (4.30, 3.66)$. Thus, strong responses to inflation and output are required to minimize the central bank's loss. I notice, however, that the loss surface is relatively flat, except when the response to output (ψ_x) is decreased below 1, which causes the loss to increase rapidly. Therefore, the loss is not very sensitive to decreases in ψ_y and ψ_x from the optimal $(\psi_y^{WG}, \psi_x^{WG})$, unless ψ_x is decreased too much.

Figure 2, plots the loss surface of the RS model; i.e., $v_{RS}(\gamma)$. The surface reaches a minimum of 11.88 at $(\psi_y^{RS}, \psi_x^{RS}) = (2.69, 1.60)$. Thus, the RS model requires smaller responses to inflation and output than the WG model, but leads to a higher loss. In fact, I find that, in contrast to the WG model, responses that are too strong to a change in inflation and output can be disastrous for the RS model, because they can lead to instability. Indeed, with responses to output bigger than 7.5, or responses to inflation bigger than 5, the RS model can become unstable. But what is striking for the RS model is that, whereas the loss surface of the WG model is relatively flat, the loss surface of the RS model is not. Therefore, the RS model is much more sensitive to changes in the Taylor rule from its optimum than the WG model is.

Figure 3 plots the loss surface of the FM model; i.e., $v_{FM}(\gamma)$. It reaches a minimum of 28.67 at $(\psi_y^{FM}, \psi_x^{FM}) = (2.39, 1.06)$. Thus, it requires milder responses to inflation and output than both the WG and RS models, but it leads to a much higher loss (about 3 times the loss of the WG model and 2.5 times the loss of the RS model). As with the RS model, the FM model is more sensitive to departures from the optimum than the WG model. The FM model is particularly sensitive to increases in the response to the output gap, but it is not as sensitive as the RS model to increases in the response to inflation.

Suppose that the central bank considers the WG, RS, and FM models to be equally plausible: $\pi_{WG} = \pi_{RS} = \pi_{FM} = 1/3$. Figures 1, 2, and 3 show that the cost of increasing $(\psi_y^{RS}, \psi_x^{RS})$ to $(\psi_y^{WG}, \psi_x^{WG})$ in the RS model, or $(\psi_y^{FM}, \psi_x^{FM})$ to $(\psi_y^{WG}, \psi_x^{WG})$ in the FM model, is much higher than the cost of decreasing $(\psi_y^{WG}, \psi_x^{WG})$ to $(\psi_y^{RS}, \psi_x^{RS})$ or $(\psi_y^{FM}, \psi_x^{FM})$ in the WG model. Indeed, $(\psi_y^{RS}, \psi_x^{RS})$ leads to a loss of 10.69 in the WG model (about 9.4 per cent higher than 9.77), while $(\psi_y^{FM}, \psi_x^{FM})$ leads to a loss of 12.43 in the WG model (about 27.2 per cent higher than 9.77). In contrast, $(\psi_y^{WG}, \psi_x^{WG})$ leads to a loss of 17.56 in the RS model (about 47.8 per cent higher than 11.88) and a loss of 43.47 in the FM model (51.6 per cent higher than 28.67). Therefore, it can be expected that the more averse the central bank is to model uncertainty, the more importance it will give to not deviating from $(\psi_y^{RS}, \psi_x^{RS})$ or $(\psi_y^{FM}, \psi_x^{FM})$. Since the FM model entails much bigger absolute minimum losses than the WG or RS model (compare 28.67 with 9.77 and 11.88), and since the central bank believes the three models to be equally plausible, it will take a central bank that is not very averse to model uncertainty to choose rules that perform very well in the WG or RS models at the expense of the FM model. Figure 4 confirms this analysis. Even when the central bank is neutral to model uncertainty, it chooses a rule $(\psi_y, \psi_x) = (2.64, 1.51)$ that already comes close to $(\psi_y^{RS}, \psi_x^{RS})$ and $(\psi_y^{FM}, \psi_x^{FM})$. As I make the central bank more model-uncertainty averse, it smoothly decreases the response to changes in inflation and output to reach $(\psi_y, \psi_x) = (2.39, 1.06)$, with an aversion coefficient of 0.5 (in the next section, I provide an interpretation of the magnitude of the aversion parameter). This draws the rule even closer to $(\psi_y^{FM}, \psi_x^{FM})$. On the other hand, if I make the central bank model-uncertainty loving, then it starts to give more importance to the WG model. It increases (ψ_y, ψ_x) gradually, and at a coefficient of model-uncertainty aversion of -1, the rule chosen is $(\psi_y, \psi_x) = (4.26, 3.63)$. This rule performs very well in the WG model, since it is very close to $(\psi_y^{WG}, \psi_x^{WG})$.

3.3.3 Optimal Taylor rules for various central bank preferences

The foregoing analysis was done for the benchmark case where the central bank assigned a weight, $\omega = 1$, to output-gap stabilization and a weight, $\nu = 0.5$, to controlling the interest rate variability. The next task is to examine how the Taylor rule behaves with respect to model-uncertainty aversion for various configurations of the central bank's preferences. Figures 5 and 6 display how ψ_y and ψ_x vary with the degree of aversion as I vary ω and ν between 0.1, 0.5, and 1. I find the panel interesting in one main respect: model uncertainty leads to quantitatively important declines in *both* the response to a change in inflation and output gap when the central bank assigns a higher weight to output stabilization relative to interest rate variability control; i.e., for the combinations $(\omega, \nu) \in \{(0.5, 0.1), (1, 0.1), (1, 0.5)\}$. This result is interesting because it shows that it is not necessary for the central bank to care too strongly about interest rate smoothing to prefer a less aggressive policy. Indeed, even when a weight of 0.1 is assigned to controlling interest rate variability, an increase in aversion to model uncertainty can lead to large declines in the coefficients of the Taylor rule (e.g., $\omega = 0.5$ and $\nu = 0.1$). In contrast, if the central bank assigns higher (or equal) weight to reducing the interest rate variability than it does to output stabilization, model uncertainty can still lead to a less aggressive policy, but often in a less important, quantitative way; e.g., $(\omega, \nu) = (0.5, 1)$.

3.3.4 Varying the priors over models

So far, I have assumed that the central bank considers all three models to be equally plausible. I now analyze how the optimal Taylor rule profiles vary as I vary the probability that the central bank assigns to the different models. Figure 7 shows how the coefficients of the Taylor rule vary with the degree of aversion to model uncertainty as I hold the weight that the central bank attaches to the RS model, π_{RS}, fixed at 0.05 and vary the weight that it assigns to the WG model, π_{WG}, between 0.05, 0.475, and 0.9. The dashed curve on the plot refers to the case where $\pi_{WG} = \pi_{FM} = 0.475$. Therefore, when the bank values the WG and FM models equally, and the RS model only slightly, the optimal Taylor rule requires more aggressive responses to changes in inflation and output when the central bank is not averse to model uncertainty, and less aggressive responses otherwise. The solid curve on the plot displays how the optimal Taylor rule profile changes when π_{WG} is increased to 0.9 and π_{RS} is held constant at 0.05 (and hence π_{FM} is decreased to 0.05). As I increase the weight given to the WG model to $\pi_{WG} = 0.9$, the path of the responses to a change in inflation and output shifts upwards. Thus, at each degree of model-uncertainty aversion, the Taylor rule requires more aggressive responses to changes in inflation and output. This happens because as the central bank believes more in the WG model, it becomes more important to choose rules which perform well in it. Hence, more aggressive rules are chosen at each degree of aversion because aggressive rules are what work best for the WG model. Conversely, if I hold π_{FM} fixed at 0.05 and decrease π_{WG} to 0.05 (and hence increase π_{RS} to 0.9), at each degree of aversion the central bank prefers rules that work better in the FM model. The profile therefore shifts downwards to the dotted curve. Figure 8, on the other hand, shows π_{FM} fixed at 0.05 and varies π_{WG} between 0.05, 0.475, and 0.9. The foregoing analysis for Figure 7 applies to this case. A higher π_{WG} relative to π_{RS} holding π_{FM} fixed at 0.05 implies that the optimal Taylor rule profile shifts upwards. Conversely, a lower π_{WG} relative to π_{RS} holding π_{FM} fixed at 0.05 shifts the profile downwards.

I also hold π_{WG} fixed and vary π_{RS} between 0.05, 0.475, and 0.9. This case differs from the previous two in that the WG model for which Taylor rules are the least restrictive (because they lead to lower minimum losses relative to the RS and FM models) is now assumed not to be very plausible. As a result, at higher levels of aversion, the central bank prefers rules that perform better in the RS and FM models; only at very low levels of aversion (in fact, negative) does the central bank care about rules that perform better in the WG model. When I increase the weight that the central bank gives to the RS model relative to the FM model, holding π_{WG} at 0.05, I find that, at the higher levels of aversion, the optimal Taylor rule profile shifts upward, as shown in Figure 9. This shift occurs because, at higher levels of aversion, the central bank cares almost exclusively about the RS and FM models. As I increase the weight to the RS model, the central bank therefore prefers slightly more aggressive rules, because the RS model requires rules that are slightly more aggressive than the FM model. On the other hand, at lower degrees of aversion, the central bank cares more about the RS and WG models, because Taylor rules are less restrictive for those two models compared with the FM model. Therefore, as I increase the weight that the bank gives to the RS model, the optimal Taylor rule profile shifts downward, reflecting the fact that the RS model requires less aggressive rules than the WG model.

4. Matching the Empirical Taylor Rule

Section 3 showed that, if the central bank assigns a high enough weight to output-gap stabilization relative to interest rate variability control, an aversion to model uncertainty can lead to relatively large declines in the coefficient of the optimal Taylor rule. Even though these declines are quantitatively important, however, the optimal Taylor rules eventually converge to rules that are still more aggressive than the empirical Taylor rule. For example, for $\omega = 1$ and $\nu = 0.5$, the optimal Taylor rule converges to $(\psi_y, \psi_x) = (2.39, 1.06)$. Therefore, while the response to the output gap is only slightly too aggressive, the response to inflation is still too aggressive. In this section, I aim to produce an optimal Taylor rule that comes closer to the empirical one. To do so, I extend the policy problem of section 3 in three alternative ways:

(i) I consider a central bank that, in addition to being uncertain across models, is also uncertain about the parameters of the models. The central bank then accounts for model and model-parameter uncertainty by choosing policy according to a criterion like (5).

(ii) I consider a central bank that, in addition to model uncertainty, also faces data uncertainty. It compares final and real-time estimates of inflation and the output gap to explicitly model their measurement-error processes. Then, taking the parameters of the measurement-error processes as given, it incorporates them into each model and accounts for data uncertainty according to a criterion like (4).

(iii) Same as item (ii) above, except that the central bank recognizes that it may not know the parameters of the measurement-error processes precisely.

4.1 Monetary policy under model and model-parameter uncertainty

Suppose that the central bank is uncertain about some of the parameters of the WG, RS, and FM models, and suppose that, to capture its uncertainty, it assumes that each parameter is independently drawn from a normal distribution, with the mean and standard error given in Tables 1, 2, and 3 for the WG, RS, and FM models, respectively.[9] This implies that the joint distributions of the parameters of the WG, RS, and FM models (i.e., $\mathcal{P}_{WG}$, $\mathcal{P}_{RS}$, and $\mathcal{P}_{FM}$) are multivariate normal distributions from which I can easily sample for my computations.

Many papers (e.g., Rudebusch 2001, Estrella and Mishkin 1999) find that model-parameter uncertainty does not yield quantitatively important effects (at least, not when the central bank's loss function is quadratic). I verify whether this is true for the WG, RS, and FM models. Table 5 displays the optimal Taylor rule under no parameter uncertainty and under parameter uncertainty captured by $\mathcal{P}_{WG}$, $\mathcal{P}_{RS}$, and $\mathcal{P}_{FM}$, respectively. Indeed, I find that parameter uncertainty has a negligible effect on the optimal Taylor rule; it is slightly less aggressive in the WG model and slightly more aggressive in the RS and FM models. Therefore, these results are consistent with the literature.

Suppose that, to account for its model and model-parameter uncertainty, the central bank chooses the coefficients, γ, of the Taylor rule by minimizing

$$\pi_{WG}\, E_{\mathcal{P}_{WG}} \phi(v^{+}_{WG}(\gamma, \theta_{WG})) + \pi_{RS}\, E_{\mathcal{P}_{RS}} \phi(v^{+}_{RS}(\gamma, \theta_{RS})) + \pi_{FM}\, E_{\mathcal{P}_{FM}} \phi(v^{+}_{FM}(\gamma, \theta_{FM})), \tag{23}$$

where $v^{+}_{WG}(\gamma, \theta_{WG})$, $v^{+}_{RS}(\gamma, \theta_{RS})$, and $v^{+}_{FM}(\gamma, \theta_{FM})$ are the indirect losses of each model given the Taylor rule (7) and the parameters θ_{WG}, θ_{RS}, and θ_{FM}, respectively. As explained in section 2.2, ϕ captures the attitude of the central bank towards both model and parameter uncertainty. I assume that ϕ is given by (22). Therefore, by increasing η, I can capture the effect of an aversion to model and parameter uncertainty on the coefficients of the Taylor rule.

Figure 10 compares how the coefficients of the Taylor rule vary with the degree of aversion when the central bank does or does not face model-parameter uncertainty for the benchmark central bank preferences $\omega = 1$ and $\nu = 0.5$, and when each model is believed to be equally plausible. The solid line corresponds to the optimal Taylor rule without parameter uncertainty, and the dotted line corresponds to the rule with parameter uncertainty. I find that, instead of making the Taylor rule less aggressive, an aversion to model and model-parameter uncertainty in fact makes it more aggressive. For lower

[9] The parameters β in the WG model and d in the FM model are not originally estimated but fixed to 0.99 and 0.9756 by the authors, since they have other evidence to support these numbers. I similarly assume no uncertainty about those parameters. Also, in the WG model, the parameters ρ_r and ρ_u are restricted to lie between 0 and 1.

degrees of aversion, the optimal Taylor rules with and without model-parameter uncertainty are relatively close to each other. In contrast, once the central bank becomes averse to model and model-parameter uncertainty ($\eta > 0$), the optimal Taylor rule with uncertainty requires more aggressive responses. This occurs because, when the central bank is averse to model *and* model-parameter uncertainty, it cares about those model-specific parameter configurations that can yield bad performances for some choices of the Taylor rule, even though they may be very unlikely. For my benchmark case, one or more parameter configurations perform badly for less aggressive Taylor rules. Therefore, as the degree of aversion is increased, these parameter configurations receive more and more importance and the central bank chooses a more aggressive Taylor rule to protect itself against them.

This result shows that, for the three models under consideration, an aversion to different model-parameter configurations is unlikely to explain why optimal Taylor rules are more aggressive than the empirically observed Taylor rule. Indeed, it in fact leads to more aggressive Taylor rules. I therefore assume in the next section that the model parameters are known by the central bank, and analyze the effect of data uncertainty on policy.

4.2 Monetary policy under model and data-parameter uncertainty

I extend the problem in section 3 to one where the central bank faces not only model uncertainty but also data uncertainty, and it wants to account for it when choosing policy. Specifically, I consider a central bank that is aware that, at any particular time, output-gap and inflation data are not flawless since they are subject to multiple revisions over time. Orphanides (2001), for instance, constructs a real-time data set of the output gap and inflation which, upon comparison with ex-post revised data, suggests that the real-time uncertainty that policy-makers face can sometimes be substantial. He further concludes that "The presence of noise in real-time estimates of inflation and the output gap must be accounted for in evaluating rules-setting policy in reaction to these variables" (page 983).

The central bank of this section takes this advice seriously by choosing policy in the form of a real-time Taylor rule,

$$i_t = \psi_y \bar{y}_{t|t} + \psi_x x_{t|t}, \tag{24}$$

where $\bar{y}_{t|t} = y_{t|t} + y_{t-1|t} + y_{t-2|t} + y_{t-3|t}$ and $x_{t|t}$ are the noisy estimates[10] of the average four-quarter inflation and output gap that the central bank has in hand at time t:

$$\bar{y}_{t|t} = \bar{y}_t + \xi_t^y, \tag{25}$$
$$x_{t|t} = x_t + \xi_t^x, \tag{26}$$

where ξ_t^y and ξ_t^x are the contemporaneous measurement errors that are uncorrelated with the true inflation, $\bar{y}_t$, and output gap, x_t, series, respectively.

What are reasonable models for the measurement errors ξ_t^y and ξ_t^x? Should I, for instance, simply assume that these are mean zero i.i.d. errors that show up quarter by quarter, or should I model them in a more sophisticated way? To answer this question, I use Orphanides' data set, which provides real-time Greenbook forecasts of the average inflation and output gap ending in 1995.[11] I assume that the Greenbook forecasts are proxies for the noisy real-time estimates of output gap and inflation, and use the 2001 Bureau of Economic Analysis data as the true estimates of output gap and inflation.[12] An immediate question that arises with this strategy is as follows: Why are the Greenbook forecasts not the best estimates of the true values, $\bar{y}_t$ and x_t? Indeed, if the Greenbook estimates are the best estimates that the central bank arrived at in real time, then ξ_t^y and ξ_t^x would be forecast errors orthogonal to the Greenbook estimates, rather than the true values. Thus, the Greenbook estimates would be invalid

[10] $z_{t-k|t}$ is the estimate of the time $t-k$ value of z at time t.

[11] I thank Athanasios Orphanides for generously providing the data.

[12] I compute the "true" output gap as the deviation of real GDP from the HP-filtered real GDP.

noisy indicators. In contrast, if the Greenbook estimates are noisy estimates of the true values, then ξ_t^y and ξ_t^x would be noise uncorrelated with the true values. An important empirical question with respect to using the Greenbook forecasts is whether the errors ξ_t^y and ξ_t^x are forecast errors or noise.

Table 7 shows the correlations[13] between the real-time output-gap data and two alternative measures of the true output gap: one based on measuring potential output by the Hodrick-Prescott (HP) filter and another using the Congressional Budget Office (CBO) (2001) estimates of potential output (see Arnold, Dennis, and Peterson 2001).[14] In both cases, I find that the errors (the difference between the real-time data and final data) are much more correlated with the real-time estimates than the final estimates (a correlation of 0.9 and 0.6, respectively, for the HP filter and 0.7 and 0.3 for CBO). In Table 8, I run a regression of the errors on the real-time data and final data: for the HP filter output gap, the real-time data explain 81 per cent of the variation in the measurement errors, whereas the final data explain 38 per cent of that variation. Moreover, the real-time data are much more statistically significant in explaining the errors than are the final data. When I compute the final output gap using CBO's potential output, the regressions show a similar pattern: the real-time data explain 50 per cent of the variation in the measurement errors, and the final data explain only 8 per cent of the variation. The real-time data are also more significant in explaining the errors than are the final data.

Similarly, for average four-quarter inflation, I find that the real-time average inflation is statistically significant, and that the final average inflation is not. The real-time data explain 5 per cent of the variation in the measurement errors, but the final data explain only 0.04 per cent of the variation.

These results seem to indicate that the measurement errors in inflation and the output gap are closer to being noise than to being forecast errors. Why the Greenbook forecasts do not seem to be the best forecasts is an interesting question that I do not address in this paper. Instead, in view of the correlations in the data, I use the Greenbook estimates as noisy indicators of inflation and the output gap, as in Rudebusch (2001) and Orphanides (2003a,b).

Figures 12 and 13 plot the real-time Greenbook estimates of the output gap and average four-quarter inflation together with their final estimates based on 2001 data. Figure 12 shows that the measurement errors in the output gap appear to be quite persistent (see 1987–88, 1991–95). I follow Rudebusch (2001) and assume that the central bank models the contemporaneous measurement error between the real-time and final output-gap series as an AR(1) process,

$$\xi_t^x = \lambda \xi_{t-1}^x + \zeta_t^x, \tag{27}$$

where ζ_t^x has the standard error $\sigma_{\zeta,x}$. I estimate the AR(1) process using the real-time and final quarterly output-gap data shown in Figure 12. I obtain $\lambda = 0.75$, and the AR(1) process explains about 55 per cent of the variation in the series.

The measurement error in inflation can be rewritten as

$$\xi_t^y = (y_{t|t} - y_t) + (y_{t-1|t} - y_{t-1}) + (y_{t-2|t} - y_{t-2}) + (y_{t-3|t} - y_{t-3}).$$

Therefore, ξ_t^y, consists of the error in predicting y_t and also revision errors that arise because my time t estimate of time $t-k$ quarterly inflation differs from the final values. Assuming that

$$y_{t-k|t} - y_{t-k} = \varphi_k (y_{t-k|t-k} - y_{t-k}),$$

[13] I take this approach following Rudebusch (2001). I obtain the same conclusions as he does, but I reproduce the tables for completeness.

[14] The CBO method starts with the framework of a Solow growth model, with a neoclassical production function at its core, and estimates trends in the components of GDP using Okun's law. According to Okun's law, actual output exceeds its potential level when the rate of unemployment is below the natural rate of unemployment. Conversely, when the unemployment rate exceeds its natural rate, output falls short of potential. Unfortunately, CBO estimates are only on an annual basis. They provide quarterly estimates that are a mechanical interpolation of the annual data. For that reason, I prefer the HP filter measure of the output gap. The two series are similar except in levels.

i.e., the revision error in the time t estimate of $t-k$ inflation is proportional to the $t-k$ forecast error of $t-k$ inflation (which is reasonable if $\varphi_k < 1$, since at time t there is more information than at time $t-k$), the measurement error between real-time and final four-quarter inflation becomes an MA(3) process:

$$\xi_t^y = \zeta_t^y + \varphi_1 \zeta_{t-1}^y + \varphi_2 \zeta_{t-2}^y + \varphi_3 \zeta_{t-3}^y, \tag{28}$$

and ζ_t^y is assumed to have standard error $\sigma_{\zeta,y}$. This is consistent with Rudebusch (2001), and the estimated values of the parameters in (28) are very similar to his estimates (see Table 6).

Sections 4.2.1 and 4.2.2 consider the effect of an aversion for model uncertainty on the coefficients of the Taylor rule for two subcases: first, when the central bank takes as given the parameters of the measurement-error processes (data uncertainty) and, second, when the central bank is unsure about the parameters $\theta = (\lambda, \varphi_1, \varphi_2, \varphi_3, \sigma_{\zeta,y}, \sigma_{\zeta,x})$, of the measurement-error processes (data-parameter uncertainty).

4.2.1 Monetary policy under model uncertainty when the measurement-error processes are known

Suppose that the central bank is confident that the structure of the measurement-error processes is correct and that the parameters $\theta = (\lambda, \varphi_1, \varphi_2, \varphi_3, \sigma_{\zeta,y}, \sigma_{\zeta,x})$ are also known. To account for its model uncertainty and real-time data uncertainty, the bank chooses policy according to

$$\pi_{WG}\, \phi(v_{WG}^*(\gamma)) + \pi_{RS}\, \phi(v_{RS}^*(\gamma)) + \pi_{FM}\, \phi(v_{FM}^*(\gamma)), \tag{29}$$

with ϕ given by (22). v_{RS}^* is the indirect loss arrived at by applying the real-time Taylor rule (24) to the RS model ((12), (13)), the relationship between real-time and final inflation and the output gap ((25), (26)) and the measurement-error processes ((28), (27)). I obtain v_{WG}^* and v_{FM}^* by similarly applying the real-time Taylor rule to the equations of the WG and FM models, respectively, the link between real-time and final estimates and the measurement-error processes.

In the analysis that follows, I take the parameters of the WG, RS, and FM models as given in Tables 1, 2, and 3 and the parameters of the measurement-error processes as given in Table 6. The parameters of the measurement-error processes are those reported in Rudebusch (2001).

Figure 11 reports how the coefficients of the Taylor rule vary with the degree of aversion to model uncertainty, as when the parameters of the measurement-error processes are known for the benchmark case, $\omega = 1$ and $\nu = 0.5$. The solid curve on the graph shows the profile of the optimal Taylor rule when there is no measurement error, and the dotted curve shows the same profile under data uncertainty. The figure shows that, for the higher degrees of aversion to model uncertainty, taking data uncertainty into account dampens the coefficients of the Taylor rule.[15] Indeed, the Taylor rule converges to $(\psi_y, \psi_x) = (2.20, 0.95)$ after a degree of aversion of about 0.2 has been reached. This compares to $(2.39, 1.06)$ without data uncertainty. Therefore, accounting for data uncertainty brings the response to a change in the output gap within the range of empirically measured Taylor rules, but it leaves the response to a change in inflation more aggressive than what is empirically observed.

4.2.2 Monetary policy under model uncertainty when the parameters of the measurement-error processes are not known precisely

In this section, I assume that although the central bank is confident that the structure of the measurement-error processes is correct, it is uncertain about the parameters, $\theta = (\lambda, \varphi_1, \varphi_2, \varphi_3, \sigma_{\zeta,y}, \sigma_{\zeta,x})$,

[15] At smaller degrees of aversion, the optimal Taylor rule under data uncertainty requires more aggressive behaviour than the optimal rule without data uncertainty, because when there is data uncertainty, the WG model requires even more aggressive Taylor rules for optimality. Since, under data uncertainty, Taylor rules still lead to lower losses in the WG model than in the RS or FM models, as I decrease the aversion of the central bank to model uncertainty, the bank chooses rules that perform better in the WG model. Hence, it chooses rules that are more aggressive than before.

of the processes. As a result, the central bank assumes that θ comes from a prior, $\mathcal{P}$ on Θ, and chooses γ by minimizing

$$\pi_{WG}\, E_{\mathcal{P}}(\, \phi(v_{WG}(\gamma,\theta))\,) + \pi_{RS}\, E_{\mathcal{P}}(\, \phi(v_{RS}(\gamma,\theta))\,) + \pi_{FM}\, E_{\mathcal{P}}(\, \phi(v_{FM}(\gamma,\theta))\,), \tag{30}$$

where $v_{WG}(\gamma,\theta)$, $v_{RS}(\gamma,\theta)$, and $v_{FM}(\gamma,\theta)$ are the indirect losses of each model, given the Taylor rule (24) and the parameters, θ, of the measurement-error processes. As explained in section 2.2, ϕ captures the attitude of the central bank towards both model and data-parameter uncertainty. I assume that ϕ is given by (22) and that, by increasing η, I can capture the effect of an aversion to model and data-parameter uncertainty on the coefficients of the Taylor rule.

4.2.2.1 Computation details

To solve for the optimal Taylor rule in (30), the prior $\mathcal{P}$ must be specified. For computational simplicity, I construct $\mathcal{P}$ as follows: I assume that each parameter, say θ_i, can take only one of three values of $\Theta_i = \{\theta_i^L, \theta_i^0, \theta_i^H : \theta_i^L < \theta_i^0 < \theta_i^H\}$ with probability $\mathcal{P}_i = \{p_i^L, p_i^0, p_i^H : p_i^L + p_i^0 + p_i^H = 1\}$. Since $\theta = (\lambda, \varphi_1, \varphi_2, \varphi_3, \sigma_{\zeta,y}, \sigma_{\zeta,x})$ consists of six parameters, this implies that the parameter space Θ consists of 3^6 six-tuples: $\theta^k = (\lambda^k, \varphi_1^k, \varphi_2^k, \varphi_3^k, \sigma_{\zeta,y}^k, \sigma_{\zeta,x}^k)$. Therefore, $\Theta = \{\theta^1, ..., \theta^k, ..., \theta^{729}\}$ and, by assuming that each parameter is independently distributed according to the binomial distribution above, the joint distribution $\mathcal{P}$ over Θ is $\mathcal{P} = \{q_1, ..., q_k, ..., q_{729} : \sum_{k=1}^{729} q_k = 1\}$, where $q_k = \prod_{i=1}^{6} p_i^{k_i}$ and $k_i \in \{L, 0, H\}$.

In the computational exercise that follows, for the parameters $\lambda, \varphi_1, \varphi_2$, and φ_3, I assume that θ_i^0 is simply the estimate of the parameter as given in Rudebusch (2001), while θ_i^L, θ_i^H are given by one standard deviation below and above the estimate, respectively. For the standard errors $\sigma_{\zeta,y}, \sigma_{\zeta,x}$, I similarly assume that θ_i^0 is the estimate reported in Rudebusch (2001), but θ_i^L, θ_i^H are, respectively, taken as 1/2 and 2 times the estimate. I also assume that $\{p_i^L, p_i^0, p_i^H\} = \{0.25, 0.5, 0.25\}$ for each parameter. Therefore, there is a greater probability that a parameter takes its estimated value, but, nevertheless, there is a non-trivial probability that it takes a low or high value.

4.2.2.2 Result

The dashed curve in Figure 11 shows how the coefficients of the Taylor rule vary with the degree of aversion to model and data-parameter uncertainty when the central bank has only the prior $\mathcal{P}$ for the parameters of the measurement-error processes, for the benchmark case $\omega = 1$ and $\nu = 0.5$. The figure shows that, when the central bank is worried about *both* model and data-parameter uncertainty, increasing the degree of aversion brings an additional attenuation of both the responses to inflation and the output gap. Indeed, relative to the case where the central bank assumes the measurement-error processes to be completely known, I find that the Taylor rule converges to $(\psi_y, \psi_x) = (1.97, 0.80)$, as opposed to $(2.20, 0.95)$. Interestingly, the additional attenuation in the response to a change in inflation brings the response within the range of empirically measured Taylor rules. This is not the case when the central bank assumes that the parameters of the measurement-error processes are known.

5. Interpreting the Degree of Aversion to Model and Data-Parameter Uncertainty

So far, I have shown that, after a certain degree of aversion towards model and data-parameter uncertainty has been reached, the coefficients of the optimal Taylor rule fall within the ranges of empirically observed Taylor rules. This prompts the following question: Is the degree of aversion at which the optimal Taylor rule matches its empirical counterpart small enough to be economically sensible?

In risk theory, to interpret the size of the risk-aversion parameter, it is related to the risk premium that the decision maker would be prepared to pay to be indifferent between accepting a small risk or obtaining a non-random amount for sure. In analogy to this, I will define ρ as the *proportional premium*[16] that the central bank would be ready to pay to be indifferent between facing model and data-parameter uncertainty or achieving the average loss for sure, where the average is taken with respect to the prior over models and parameters:

$$E_\pi \phi(\tilde{v}) = \phi((E_\pi \tilde{v})(1+\rho)), \tag{31}$$

where $\tilde{v}$ is the loss of a particular model and π is the joint distribution over models and parameters. In the neighbourhood of $E_\pi \tilde{v}$, a second-order expansion of the left-hand side of (31) and a first-order expansion of its right-hand side yield

$$E_\pi \phi(\tilde{v}) \approx \phi(E_\pi \tilde{v}) + \frac{1}{2}\phi''(E_\pi \tilde{v}).\text{var}_\pi \tilde{v}, \tag{32}$$

$$\phi((E_\pi \tilde{v})(1+\rho)) \approx \phi(E_\pi \tilde{v}) + \phi'(E_\pi \tilde{v}).E_\pi \tilde{v}.\rho, \tag{33}$$

where $\text{var}_\pi \tilde{v} = E_\pi(\tilde{v} - E_\pi \tilde{v})^2$ is the variance of the models' losses given the joint distribution of models and parameters, π. Equating (32) and (33) and using the fact that $\eta = \frac{\phi''}{\phi'}$, I obtain

$$\rho = \frac{1}{2}.\frac{\text{var}_\pi \tilde{v}}{E_\pi \tilde{v}}.\eta. \tag{34}$$

Therefore, in the neighbourhood of the average loss, the proportion of the average loss that the central bank would accept to be indifferent between not facing and facing model and data-parameter uncertainty is equal to half the degree of uncertainty aversion times the variance of the models' losses per expected unit of loss.

To gauge the economic significance of this premium, it is useful to consider historical variations in the value of the loss function in the United States. Table 11 shows the value of the central bank's loss function[17] under the tenures of Arthur Burns (1970Q1–1978Q1), Paul Volcker (1979Q3–1987Q2), and Alan Greenspan (1987Q3–present). Under the (strong) assumption that these Fed chairmen had the same preferences (same ω and ν), I find that the value of the loss function under Burns and Volcker was, respectively, about 2 to 4 and 3 to 4 times the loss under Greenspan.[18] Therefore, if each regime is interpreted as one model and the loss under Greenspan is taken as the average loss around which the central bank is asked to quote the premium it would pay to avoid model and data-parameter uncertainty, a premium of even up to 100 per cent would appear to be reasonable.

Suppose that the central bank is willing to pay a premium of 30 per cent of the average loss. One can ask: What degree of model-uncertainty aversion corresponds to a 30 per cent premium? To be able to answer this, one must know the ratio $\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})}$. Table 10 calculates this ratio under the empirically observed Taylor rule[19] over 1987Q3 to 2003Q2, $i_t = 1.90\bar{y}_t + 0.78x_t$. Suppose additionally that the central bank assigns weights $\omega = 1$ and $\nu = 0.5$ to output stabilization and interest rate variability control, and assume that it believes much more in the WG model ($\pi_{WG} = 0.90, \pi_{RS} = \pi_{FM} = 0.05$). Then, $\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})} = 0.68$. This implies that the degree of model-uncertainty aversion that yields a premium

[16] I focus on the proportional premium rather than the *absolute premium* because this is more meaningful when comparing the estimate with the data.

[17] I use the sample variance of inflation, output gap, and change in interest rate to approximate $\text{var}(y_t) + \omega\text{var}(x_t) + \nu\text{var}(i_t - i_{t-1})$.

[18] Even if they are not assumed to have the same preferences, the losses under Burns and Volcker are, respectively, about 1.25 and 2.5 times greater than the loss under Greenspan (the minimum losses under Burns and Volcker (respectively, 3.19 and 5.29) arise under the preferences $\omega = 0.1$ and $\nu = 1$, whereas the maximum loss under Greenspan (1.85) arises under $\omega = 1$ and $\nu = 0.1$).

[19] I obtain this by running an OLS regression of the quarterly federal funds rate on average inflation over the past four quarters, and on the output gap, as measured by the CBO.

of 30 per cent is 0.885. With $\eta = 0.89$, the optimal Taylor rule has coefficients $(\psi_y, \psi_x) = (1.99, 0.81)$. In contrast, the optimal Taylor rule when $\eta = 0$ is $(\psi_y, \psi_x) = (3.25, 1.99)$. Thus, a model-uncertainty aversion of 0.89 yields a modest premium, but relatively large declines in the responses to changes in inflation and the output gap if the central bank believes much more in the WG model than in the RS and FM models. In contrast, if the central bank gives equal weight to all models, a premium of 30 per cent implies $\eta = 0.27$. This yields a Taylor rule with coefficients $(\psi_y, \psi_x) = (2.17, 1.03)$. The optimal Taylor rule when $\eta = 0$ is $(\psi_y, \psi_x) = (2.34, 1.18)$. Thus, the declines in the coefficients of the Taylor rule are quite small.

These results are not surprising. Figures 8 and 4 show that, when a high weight is given to the WG model relative to the other two models, the optimal Taylor rule calls for much stronger responses to inflation and the output gap when $\eta = 0$ than when all the models are regarded to be equally plausible. This occurs because, when the central bank believes much more in the WG model and is neutral to model uncertainty, it chooses rules that perform better in the WG model than in the RS or FM models. Responses to inflation and the output gap that are too strong, however, can lead the losses in the RS and FM models to increase quite rapidly. Therefore, as I make the central bank more averse to model uncertainty, even a small degree of aversion causes the central bank to be very concerned about avoiding sharp increases in the losses of the RS and FM models. In contrast, when all the models are equally weighted, at the outset, the central bank chooses a rule that performs well in the RS and FM models when $\eta = 0$. Hence, increasing the aversion to model uncertainty does not have a very large effect, because the optimal rule is already close to what is optimal in the RS and FM models.

Alternatively, I can calculate the premium that the central bank would pay for a given degree of model-uncertainty aversion. I take the same central bank preferences $(\omega = 1, \nu = 0.5)$ and assume that the central bank's degree of aversion is $\eta = 1$. If the central bank gives a weight $\pi_{WG} = 0.90, \pi_{RS} = \pi_{FM} = 0.05$ to the WG, RS, and FM models, respectively, $\eta = 1$ corresponds to a premium of 33.9 per cent and an optimal Taylor rule $(\psi_y, \psi_x) = (1.97, 0.80)$. Therefore, when the central bank believes much more in the WG model than in the RS and FM models, a degree of aversion equal to 1 leads to large declines in the coefficients of the Taylor rule relative to the case where the bank minimizes a Bayesian criterion (i.e., compare $(\psi_y, \psi_x) = (1.97, 0.80)$ with $(\psi_y, \psi_x) = (3.25, 1.99)$ when $\eta = 0$). More importantly, $\eta = 1$ generates an optimal Taylor rule that matches the empirical Taylor rule. Therefore, since $\eta = 1$ corresponds to a premium of only 33.9 per cent, under the assumption that the central bank gives a much higher weight to the WG model than to the RS or FM models, it is economically sensible to claim that an aversion to model and data-parameter uncertainty is at least partly the reason why the central bank may, in practice, prefer less aggressive Taylor rules.

On the other hand, if $\pi_{WG} = \pi_{RS} = \pi_{FM} = 1/3$, $\eta = 1$ yields a premium of 109.23 per cent and an optimal Taylor rule of $(\psi_y, \psi_x) = (1.97, 0.80)$. Therefore, when all models are weighted equally, a degree of aversion, $\eta = 1$, can generate an optimal Taylor rule that matches the empirical Taylor rule. However, since $\eta = 1$ corresponds to quite a large premium, if the central bank values all three models equally, an aversion to model and data-parameter uncertainty is a less economically plausible explanation for the attenuation of the coefficients of the Taylor rules in practice.

6. Relation to the Literature

Rudebusch (2001) investigates how much and what type of uncertainty can help to reconcile the optimal and empirical Taylor rule. Considering each of parameter uncertainty, model perturbation,[20] and data uncertainty in isolation, he finds that parameter uncertainty is irrelevant, while reasonable degrees of

[20]Rudebusch defines model perturbation as the case where the parameters of the model differ from their estimates *even on average*. This differs from parameter uncertainty in that, with parameter uncertainty, parameters are assumed to be distributed around their estimates.

data uncertainty do not lead to large enough declines in the coefficient of the optimal Taylor rule. My result that data uncertainty on its own is not enough to match the empirical Taylor rule, in the context of the decision-making problem examined in this paper, is in line with Rudebusch's finding.

Rudebusch combines model perturbation together with data uncertainty. He considers different parameter values for the interest rate sensitivity of output and the mean reversion of inflation while allowing for an empirically plausible degree of measurement error in the output gap and inflation. He finds that if the interest rate sensitivity of output is higher than estimated and/or inflation mean-reverts more slowly than estimated, data uncertainty can generate an optimal Taylor rule in line with the empirical Taylor rule. Although Rudebusch succeeds in matching the empirical rule, however, he does not answer the important question: How does the policy-maker discriminate between different parameter configurations? And what makes the policy-maker give more weight to some configurations than to others?

In the context of this paper, a desire for robustness makes the policy-maker discriminate between various configurations. Indeed, I can view the policy-maker as one who contemplates various model and data-parameter configurations. But the policy-maker's aversion to having to choose between many configurations makes them trade off average performance for robustness. It is this quest for robustness that eventually drives the policy-maker to give more weight to a configuration that yields an optimal rule which matches the empirical one.

Levin and Williams (2003) consider a central bank that values different competing reference models.[21] Their objective is to identify a rule that is robust across these different models. They find that it is possible to identify such rules as long as the policy-maker cares strongly enough about stabilizing the output gap. It is interesting to note that the same condition that generates robust policy rules in their context is also sufficient to generate large declines in the coefficients of the optimal Taylor rule in this paper.

7. Conclusion

The first part of this paper considered a central bank that faces model uncertainty because it finds three non-nested models of the economy to be plausible: Woodford and Giannoni's forward-looking model (WG), Rudebusch and Svensson's empirical backward-looking model (RS), and Fuhrer and Moore's contracting model (FM). The central bank accounts for its model uncertainty in a framework that exhibits the non-reduction of two-stage lotteries developed by Segal (1990), Klibanoff, Marinacci, and Mukerji (2002), and Ergin and Gul (2004). This means that, for the purposes of this paper, the central bank considers model risk as a risk distinct from the first-order risk that arises because of the stochastic properties of a particular model. Similar to Klibanoff, Marinacci, and Mukerji (2002), I interpret the central bank's attitude to this second-order risk as its attitude towards model uncertainty. The central bank's policy problem is to choose a Taylor rule that works reasonably well in all models given its degree of aversion to model uncertainty. I have attempted to answer the question: Does model-uncertainty aversion make policy more or less aggressive?

Given my model set, I find that an aversion to model uncertainty indeed makes policy less aggressive. But if, in addition, the central bank assigns a higher weight to output stabilization relative to interest rate variability control, an aversion to model uncertainty generates quantitatively important declines in the coefficients of the Taylor rule. And this occurs *even* if the central bank in fact assigns little weight to controlling the variability of the interest rate in its loss function. This result is interesting because many authors have argued that one of the reasons why policy responses are smooth is that a central bank's loss function assigns a high weight to interest rate smoothing. In contrast, I find that, in the presence of model uncertainty, this is not necessary. Model uncertainty can still lead to less aggressive

[21] Their set of models includes the WG model, the RS model and Fuhrer's (2000) habit-based model.

policy when the central bank cares strongly about controlling the interest rate variability relative to output stabilization, but the decline is quantitatively less important.

Even though an aversion to model uncertainty seems to be able to generate quite important declines in the coefficients of the Taylor rule, for the various central bank preferences that I considered, it still leads to optimal Taylor rules that are more aggressive (especially in the response to inflation) than empirically observed Taylor rules. The second part of this paper considered a central bank that faces not only model uncertainty but also data uncertainty. The central bank, as a result, models the measurement-error processes that plague the inflation and output-gap data that it has in real time and chooses policy according to a real-time Taylor rule. I expanded the decision-making framework to one where the central bank could also be averse to uncertainty about the parameters of the measurement-error processes, and I analyzed the effect of aversion to model and data-parameter uncertainty on the optimal Taylor rule. I found that, when the central bank faces both model uncertainty and data-parameter uncertainty, an increase in the degree of aversion leads to an optimal Taylor rule, the coefficients of which fall within the range of empirically estimated Taylor rules.

Finally, I interpreted the economic significance of the degree of model-uncertainty aversion by relating it to the proportional premium that the central bank would pay to be indifferent between facing model uncertainty or achieving the average loss of the models for sure. I found that, when the central bank assigns a higher weight to output stabilization than to interest rate smoothing, small degrees of aversion are enough to generate relatively important declines in the responses to inflation and the output gap provided that the central bank believes much more in the WG model than in the RS and FM models. For the benchmark case, where the central bank believes much more in the WG model than in the other models, I found that a degree of aversion of 1 is enough to generate an optimal Taylor rule that matches the empirical Taylor rule. This degree of aversion was shown to be small in that it corresponds to a small premium. This result suggests that, if the central bank assigned a higher weight to the WG model than to the RS or FM models, it would be economically sensible to claim that an aversion to model and data-parameter uncertainty is at least part of the reason why the central bank may, in practice, prefer more attenuated Taylor rules.

References

Anderson, G. and G. Moore. 1985. "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models." *Economic Letters* 17(3): 247–52.

Arnold, R., R. Dennis, and J. Peterson. 2001. "CBO's Method for Estimating Potential Output: An Update." Congressional Budget Office Paper.

Brainard, W. 1967. "Uncertainty and the Effectiveness of Policy." *American Economic Review* 58: 411–25.

Calvo, G. 1983. "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics* 12: 383–98.

Currie, D. and P. Levine. 1993. *Rules, Reputation and Macroeconomic Policy Coordination*. Cambridge: Cambridge University Press.

Dennis, R. 2003. "The Policy Preference of the U.S. Federal Reserve." Federal Reserve Bank of San Francisco Working Paper.

Erceg, C. and A. Levin. 2002. "Optimal Monetary Policy with Durable and Non-Durable Goods." European Central Bank Working Paper No. 179.

Ergin, H. and F. Gul. 2004. "A Subjective Theory of Compound Lotteries." M.I.T. Working Paper.

Estrella, A. and F.S. Mishkin. 1999. "Rethinking the Role of NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty." In *Monetary Policy Rules*, edited by J.B. Taylor, 405–30. Chicago: University of Chicago Press.

Fuhrer, J.C. 2000. "Habit Formation in Consumption and Its Implications for Monetary-Policy Models." *American Economic Review* 90(3): 367–90.

Fuhrer, J.C. and G.R. Moore. 1995. "Monetary Policy Trade-Offs and the Correlation between Nominal Interest Rates and Real Output." *American Economic Review* 85(1): 219–39.

Giannoni, M.P. 2000. "Optimal Interest Rate Rules in a Forward-looking Model and Inflation Stabilization versus Price Level Stabilization." Columbia University Working Paper.

Hansen, L.P. and T.J. Sargent. 2004. "Robust Control and Model Uncertainty in Macroeconomics." University of Chicago and New York University Manuscript.

Klibanoff, P., M. Marinacci, and S. Mukerji. 2002. "A Smooth Model of Decision Making under Ambiguity." International Centre for Economic Research.

Levin, A.T. and J.C. Williams. 2003. "Robust Monetary Policy with Competing Reference Models." *Journal of Monetary Economics* 50(5): 945–75.

Onatski, A. and J.H. Stock. 2002. "Robust Monetary Policy under Model Uncertainty in a Small Model of the U.S. Economy." *Macroeconomic Dynamics* 6: 85–110.

Orphanides, A. 2001. "Monetary Policy Rules Based on Real-Time Data." *American Economic Review* 91(4): 964–85.

———. 2003a. "Monetary Policy Evaluation with Noisy Information." *Journal of Monetary Economics* 50(3): 605–31.

———. 2003b. "The Quest for Prosperity without Inflation." *Journal of Monetary Economics* 50(3): 633–63.

Peersman, G. and F. Smets. 1999. "The Taylor Rule: A Useful Monetary Policy Benchmark for the Euro Area?" *International Finance* 2(1): 85–116.

Rudebusch, G.D. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83(2): 203–17.

Rudebusch, G.D. and L.E.O. Svensson. 1999. "Policy Rules for Inflation Targeting." In *Monetary Policy Rules*, edited by J.B. Taylor, 203–46. Chicago: University of Chicago Press.

Sack, B. 2000. "Does the Fed Act Gradually? A VAR Analysis." *Journal of Monetary Economics* 46(1): 229–56.

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

Segal, U. 1990. "Two-Stage Lotteries without the Reduction Axiom." *Econometrica* 58(2): 349–77.

Soderstrom, U. 1999. "Should Central Banks Be More Aggressive?" Sveriges Riksbank Working Paper.

Soderstrom, U., P. Soderlind, and A. Vredin. 2002. "Can a Calibrated New-Keynesian Model of Monetary Policy Fit the Facts?" Sveriges Riksbank Working Paper.

Svensson, L.E.O. 2003. "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules." *Journal of Economic Literature* 41(2): 426–77.

Taylor, J.B. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88(1): 1–23.

———. 1993. "Discretion vs Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39: 195–214.

Tetlow, R.J. and P. von zur Muehlen. 2001. "Robust Monetary Policy with Misspecified Models: Does Model Uncertainty Always Call for Attenuated Policy?" *Journal of Economic Dynamics and Control* 25(6): 911–49.

Woodford, M. 1996. "Control of the Public Debt: A Requirement for Price Stability?" National Bureau of Economic Research Working Paper No. 5684.

———. 1999. "Inflation Stabilization and Welfare: The Case of a Distorted Steady State." Princeton University Working Paper.

———. 2000. "A Neo-Wicksellian Framework for the Analysis of Monetary Policy." Princeton University Working Paper.

Table 1: Parameters of Woodford and Giannoni's Model

Parameters	Estimate	s.d.
Structural		
β	0.99	
σ^{-1}	1.5913	0.3322
κ	0.0952	0.014
Shock process		
ρ_r	0.35	0.10
σ_r^2	13.8266	
ρ_u	0.35	0.10
σ_u^2	0.1665	

Table 2: Parameters of Rudebusch and Svensson's Model

Parameters	Estimate	s.d.
Structural		
b_0	0.7	0.08
b_1	-0.1	0.1
b_2	0.28	0.1
b_x	0.14	0.03
d_0	1.16	0.08
d_1	-0.25	0.08
d_i	0.1	0.03
Shock process		
$\sigma_{e_\pi}^2$	1.018	
$\sigma_{e_x}^2$	0.67	

Table 3: Parameters of Fuhrer and Moore's Model

Parameters	Estimate	s.d.
Structural		
a_1	1.340	0.094
a_2	-0.372	0.09
a_ρ	0.360	0.216
d	0.9756	
s	0.08	.013
r	0.00252	0.0012
Shock process		
$\sigma^2_{e_w}$	0.07932	
$\sigma^2_{e_x}$	1.357	

Table 4: Optimal Taylor Rules for Each Model under Different (ω, ν)'s

		WG		RS		FM	
ω	ν	ψ_y	ψ_x	ψ_y	ψ_x	ψ_y	ψ_x
0.1	0.1	10.763	1.227	4.615	2.039	4.715	0.310
0.1	0.5	7.896	0.110	3.247	1.196	3.005	0.097
0.1	1	5.654	0.037	2.801	0.926	2.490	0.066
0.5	0.1	9.017	9.652	3.830	2.341	3.692	1.446
0.5	0.5	6.452	1.382	2.910	1.425	2.629	0.664
0.5	1	8.511	0.439	2.577	1.110	2.281	0.460
1	0.1	9.227	20.296	3.395	2.531	3.272	2.167
1	0.5	4.296	3.658	2.691	1.604	2.391	1.055
1	1	5.792	1.414	2.419	1.262	2.120	0.755

Table 5: Optimal Taylor Rule With and Without Model-Parameter Uncertainty when $\omega = 1$ and $\nu = 0.5$

	without		with	
Model	ψ_y	ψ_x	ψ_y	ψ_x
WG	4.296	3.658	4.1975	3.5831
RS	2.691	1.604	2.7943	1.8686
FM	2.391	1.055	2.5862	1.1179

Table 6: Parameters of the Measurement-Error Processes

Parameters	Value
Inflation	
φ_1	0.63
φ_2	0.26
φ_3	0.18
$\sigma^2_{\zeta,y}$	0.1024
Output gap	
λ	0.75
$\sigma^2_{\zeta,x}$	0.7022

Table 7: Correlations between Measurement-Errors, Final, and Real-Time Data

Measurement errors	Final	Real time
Output gap, ξ^x_t	x_t	r_t
HP	0.6167	0.9024
CBO	0.2831	0.7102
Inflation, ξ^y_t	$\bar{y}_t$	$\bar{s}_t$
	0.0639	0.2191

Table 8: Regressing the Measurement Error on the Final and Real-Time Output Gap; Output Gap based on the HP Filter and CBO Potential Output, Respectively

$x_{t\|t} - x_t$	HP		CBO	
	Real time	Final	Real time	Final
Constant	-0.67	-1.07	0.16	-0.09
	(0.09)	(0.15)	(0.14)	(0.19)
Output gap	0.51	0.64	0.35	0.19
	(0.042)	(0.14)	(0.06)	(0.11)
R^2	0.81	0.38	0.50	0.08

Table 9: Regressing the Measurement Error on Final and Real-Time Average Inflation

$\bar{y}_{t\|t} - \bar{y}_t$	Real time	Final
Constant	0.006	0.11
	(0.092)	(0.09)
Inflation	0.03	0.01
	(0.019)	(0.02)
R^2	0.05	0.004

Table 10: Interpreting the Degree of Model-Uncertainty Aversion; expected loss of each model is evaluated under the Taylor rule, $i_t = 1.897\bar{y}_t + 0.775x_t$

ω	ν	$\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})}$ $\pi_{WG} = 0.90$ $\pi_{RS} = 0.05$	$\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})}$ $\pi_{WG} = 0.05$ $\pi_{RS} = 0.90$	$\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})}$ $\pi_{WG} = 0.05$ $\pi_{RS} = 0.05$	$\frac{\text{var}_\pi(\tilde{v})}{E_\pi(\tilde{v})}$ $\pi_{WG} = 0.33$ $\pi_{RS} = 0.33$
0.1	0.1	2.3471	0.5116	0.5413	2.6303
0.1	0.5	0.8182	0.4229	0.4401	1.6159
0.1	1	0.3185	0.458	0.4389	1.0801
0.5	0.1	1.1897	0.5529	0.5807	2.3022
0.5	0.5	0.6434	0.5533	0.5127	1.7404
0.5	1	0.3669	0.6527	0.5139	1.4367
1	0.1	0.9731	0.7707	0.6909	2.5156
1	0.5	0.6783	0.8301	0.6407	2.1846
1	1	0.4964	0.9739	0.6405	2.0156

Table 11: Empirical Loss: $\frac{1}{1+\omega+\nu}\left(\widehat{\text{var}(y_t)} + \omega\,\widehat{\text{var}(x_t)} + \nu\,\widehat{\text{var}(\Delta i_t)}\right)$

ω	ν	Burns 1970Q1-1978Q1	Volcker 1979Q3-1987Q2	Greenspan 1987Q3-
0.1	0.1	4.48	6.77	1.17
0.1	0.5	3.73	5.91	0.94
0.1	1	3.19	5.29	0.77
0.5	0.1	4.68	6.63	1.56
0.5	0.5	4.04	5.97	1.30
0.5	1	3.53	5.43	1.09
1	0.1	4.83	6.53	1.85
1	0.5	4.29	6.01	1.59
1	1	3.82	5.56	1.36

Table 12: Optimal Generalized Taylor Rule: $i_t = \rho i_{t-1} + \psi_y y_t + \psi_x x_t$

			Loss		
ρ	ψ_y	ψ_x	WG	RS	FM
1.56	2.55	2.49	7.70	∞	379.85
0.25	1.84	0.98	11.94	13.81	28.18
0.87	0.46	0.47	13.12	21.83	24.73

Figure 1: Loss surface of WG model; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$

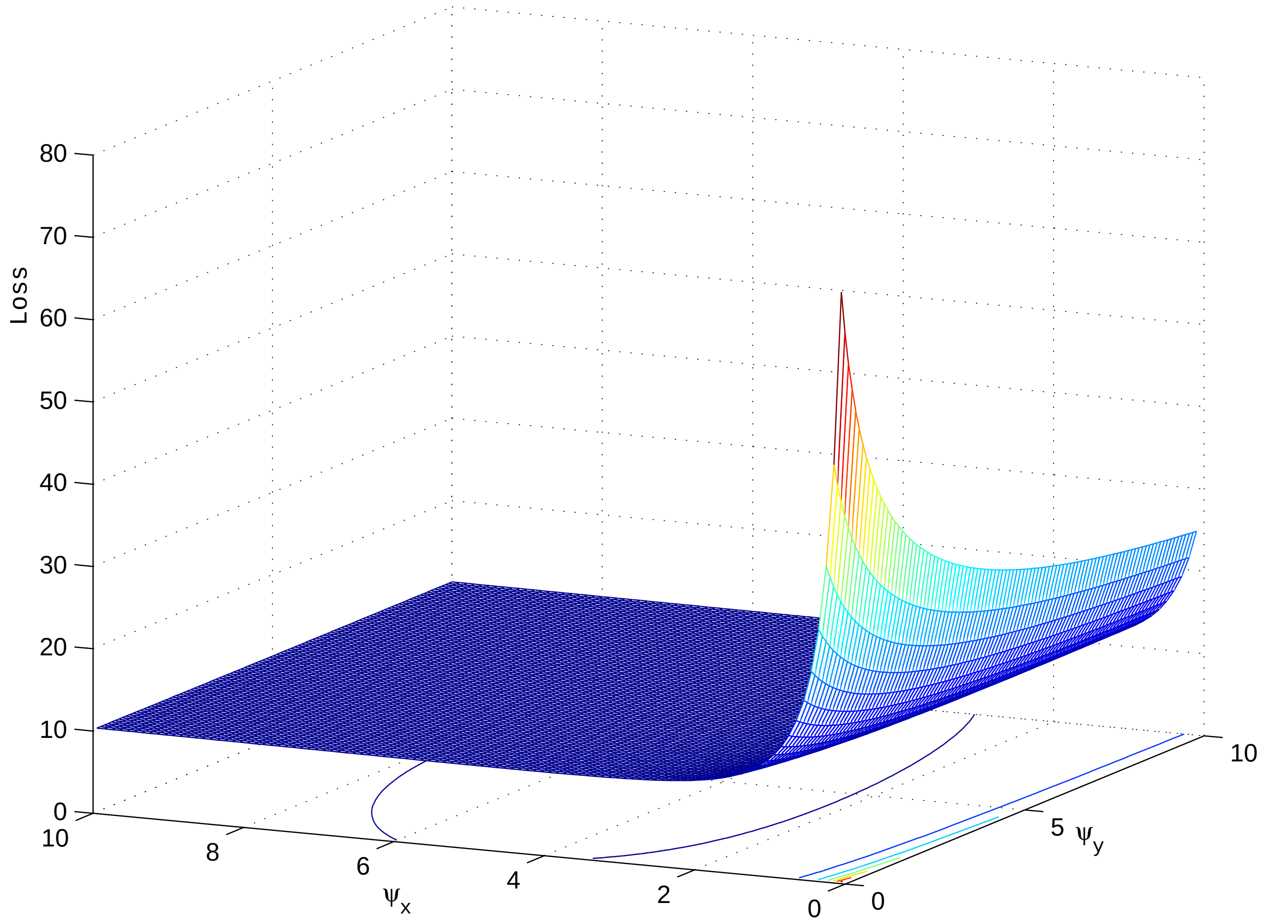


Figure 2: Loss surface of RS model; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$

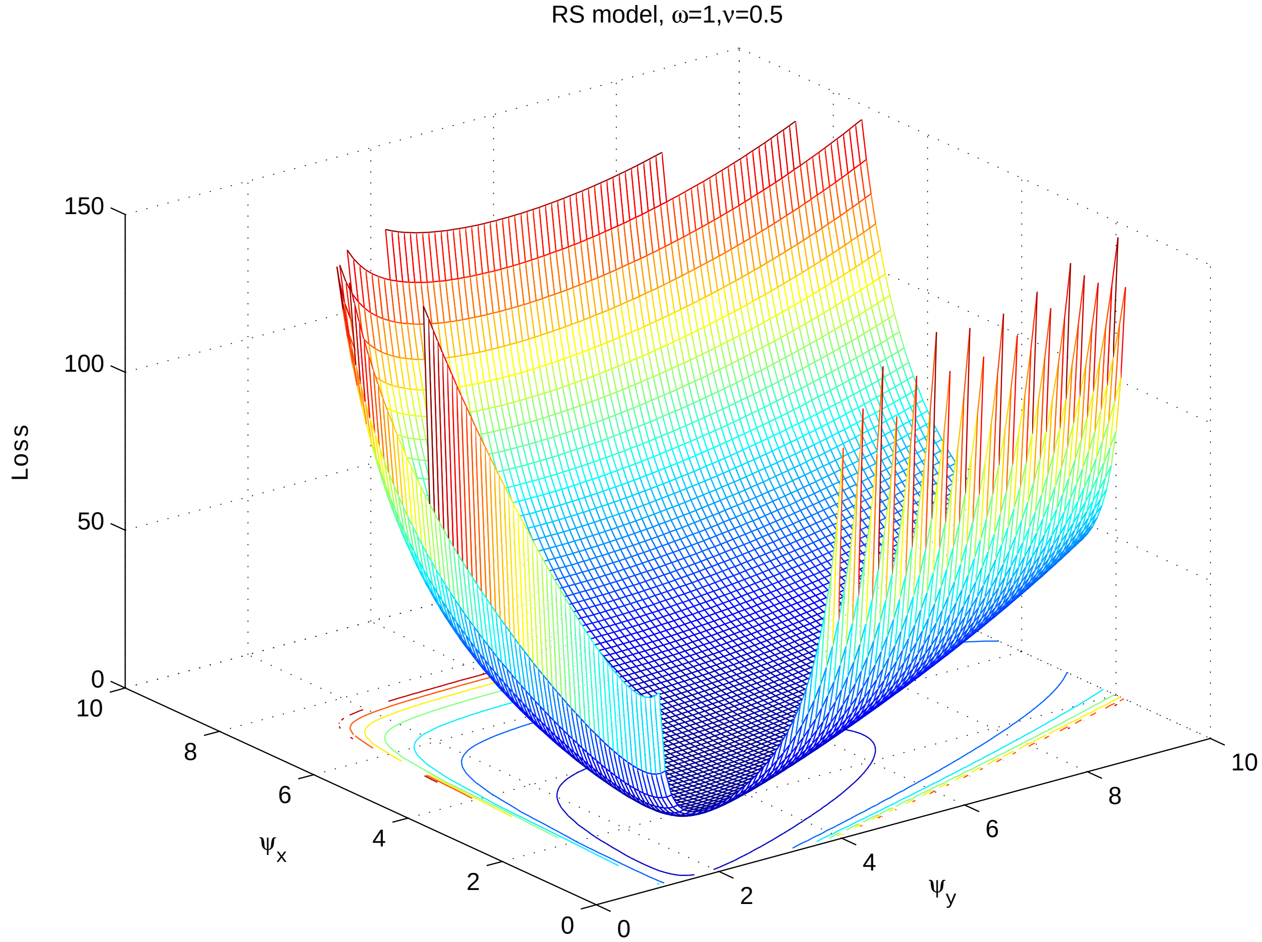


Figure 3: Loss surface of FM model; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$

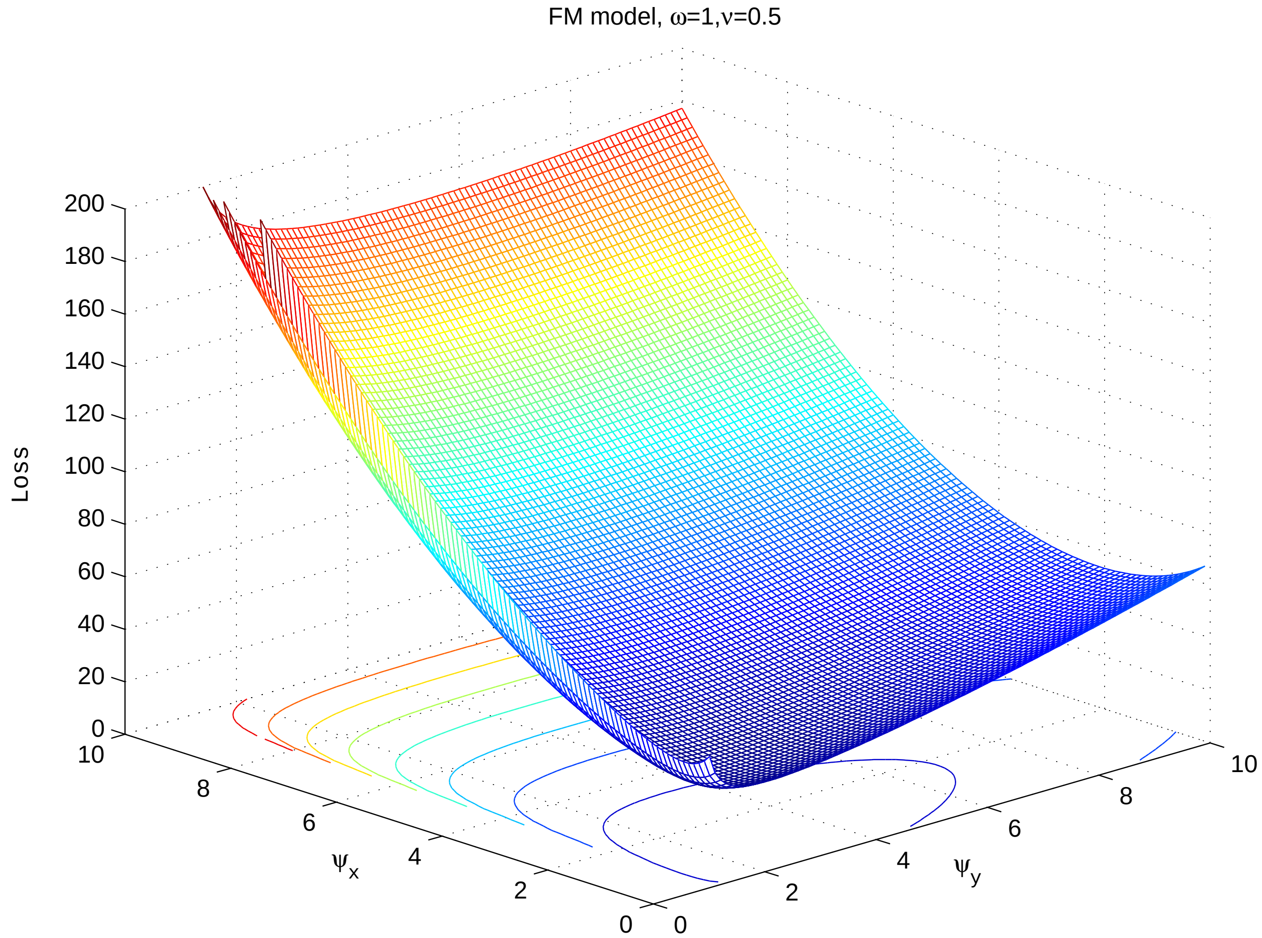


Figure 4: How the Taylor rule varies with the degree of model-uncertainty aversion; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$

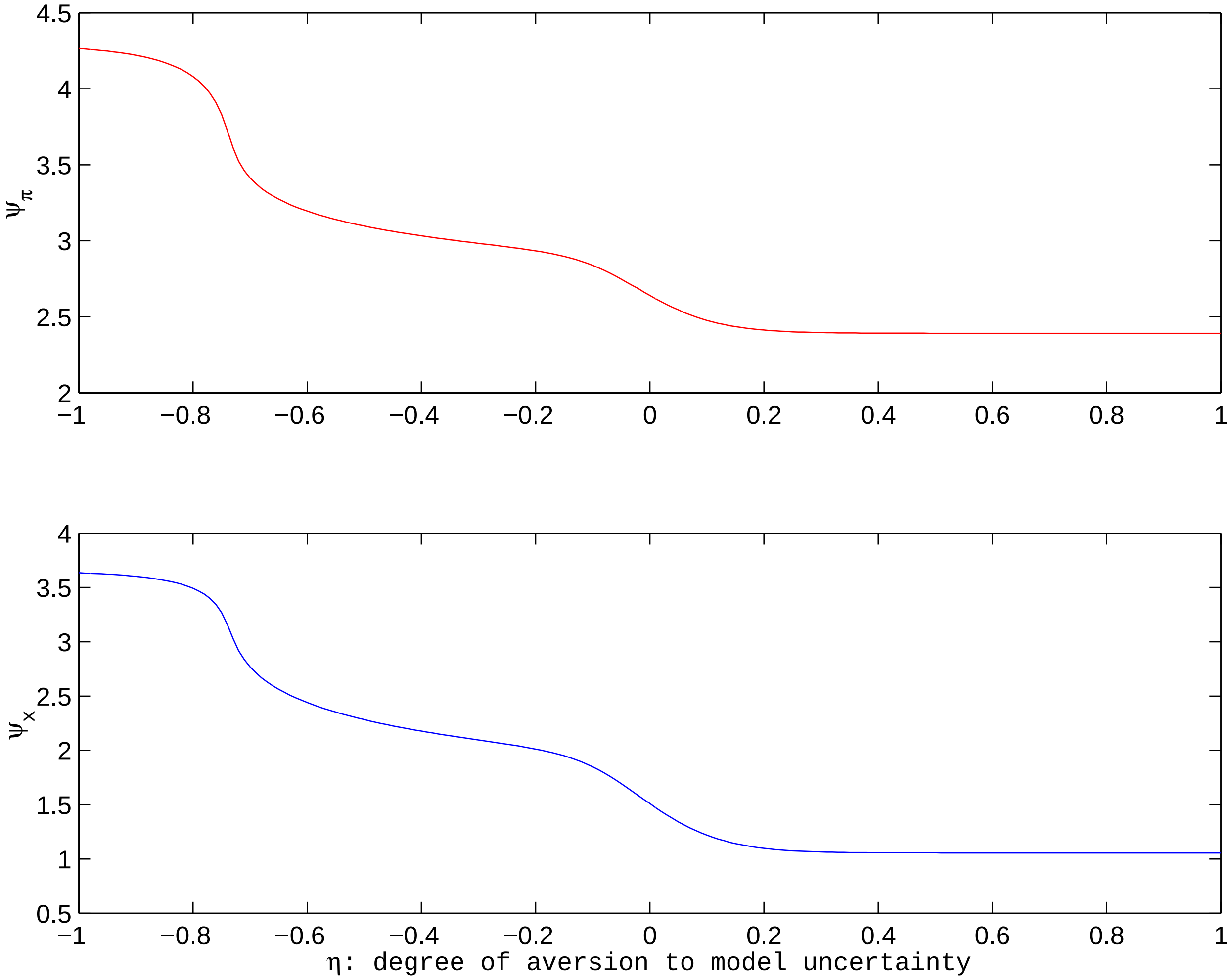


Figure 5: How the response to inflation, ψ_y, varies with model-uncertainty aversion under different central bank preferences

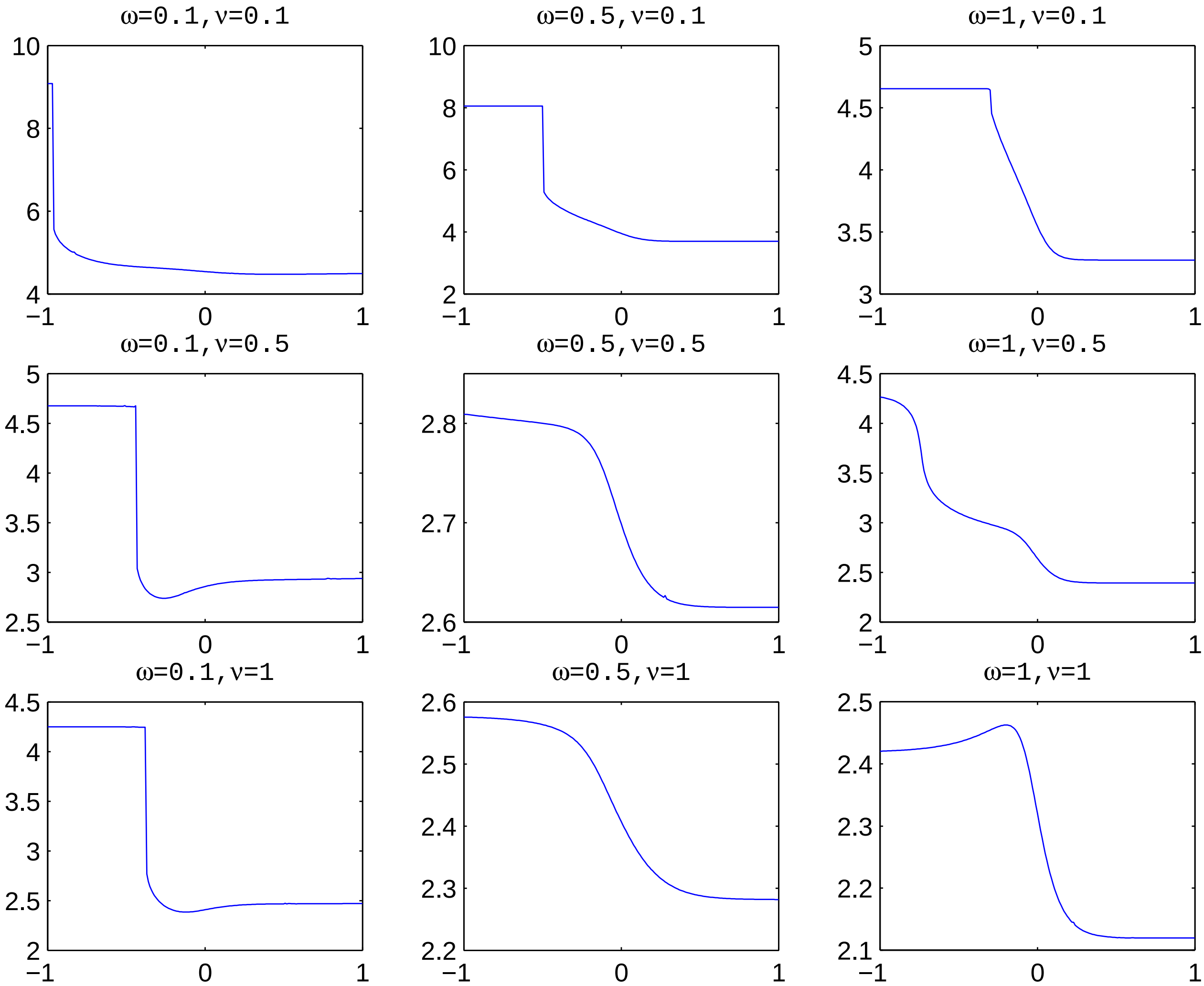


Figure 6: How the response to the output gap, ψ_x, varies with model-uncertainty aversion under different central bank preferences

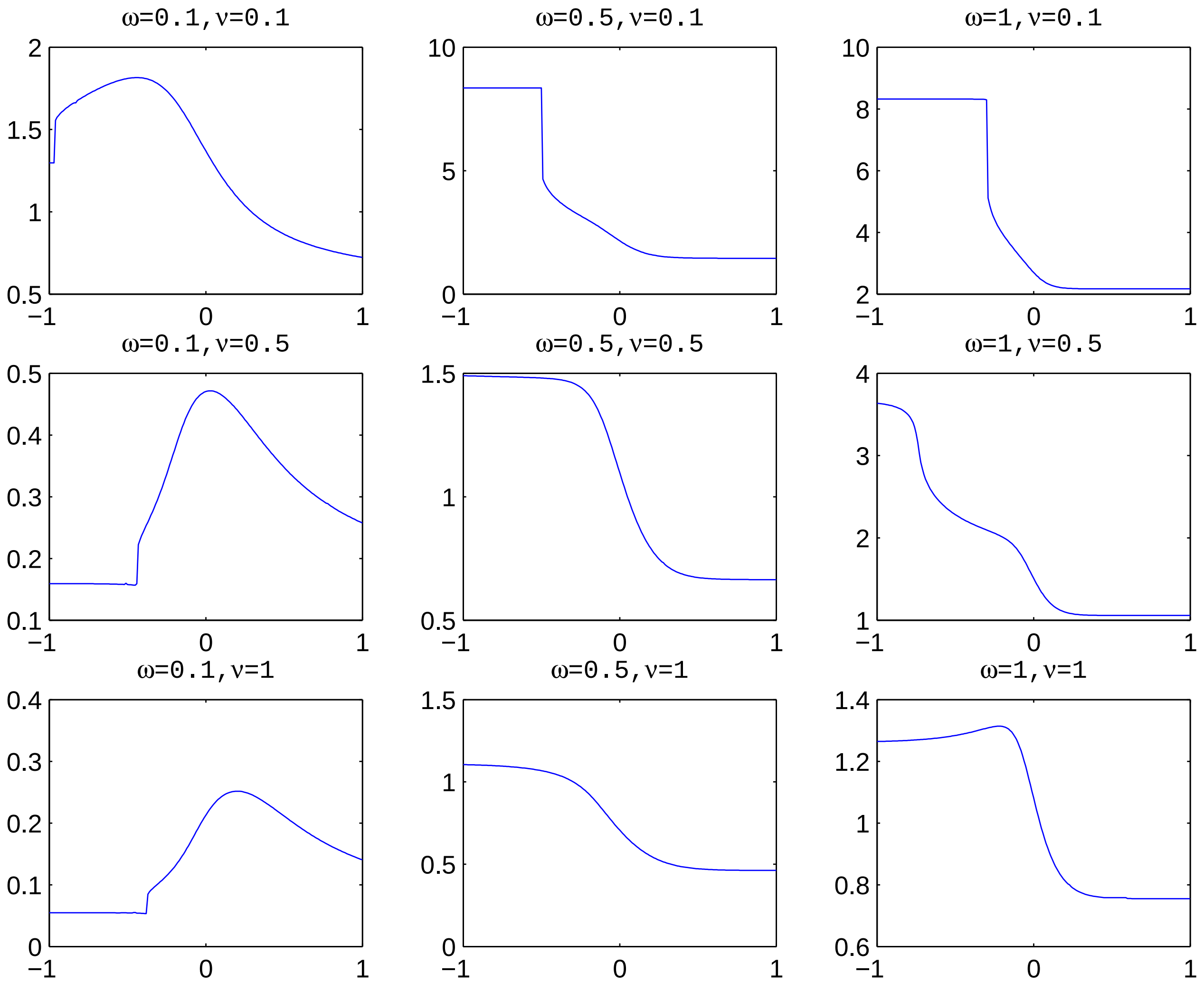


Figure 7: How the Taylor rule varies with the degree of model-uncertainty aversion; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$; weight to WG model, π_{WG}, varies between 0.05, 0.475, and 0.9; weight to RS model, π_{RS}, is held fixed at 0.05

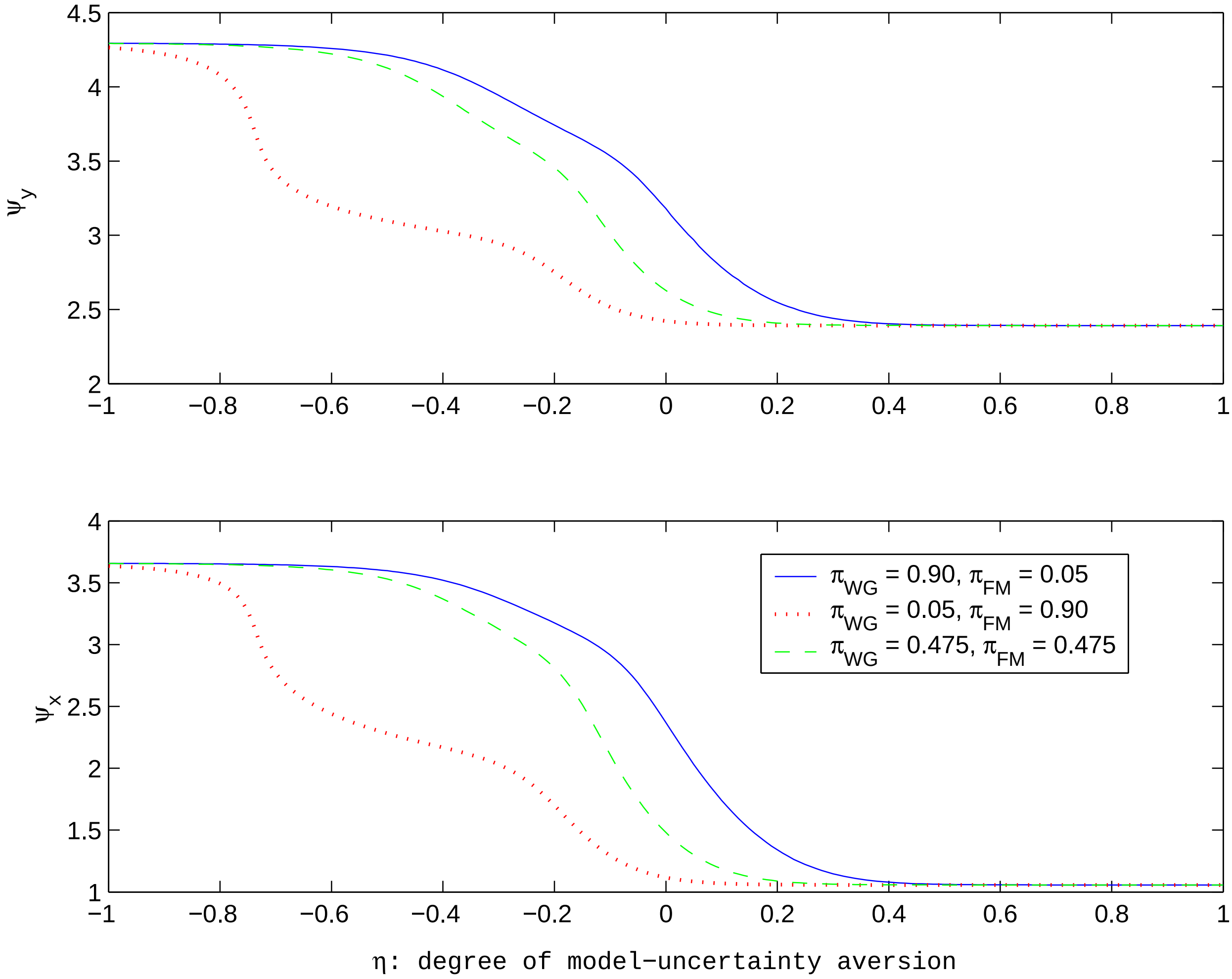


Figure 8: How the Taylor rule varies with the degree of model-uncertainty aversion; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$; weight to WG model, π_{WG}, varies between 0.05, 0.475, and 0.9; weight to FM model, π_{FM}, is held fixed at 0.05

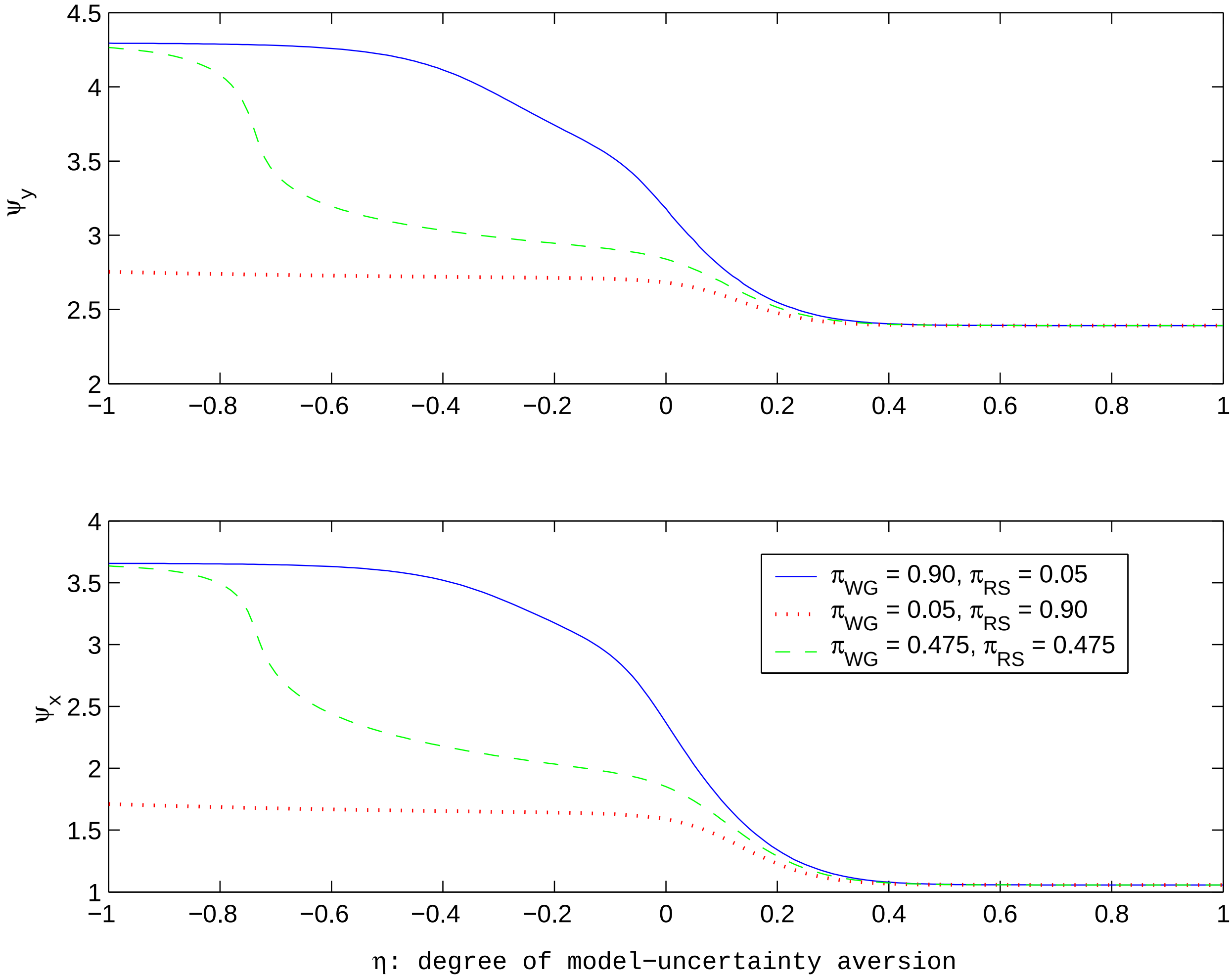


Figure 9: How the Taylor rule varies with the degree of model-uncertainty aversion; weight to output gap is $\omega = 1$ and weight to interest rate smoothing is $\nu = 0.5$; weight to RS model, π_{RS}, varies between 0.05, 0.475, and 0.9; weight to WG model, π_{WG}, is held fixed at 0.05

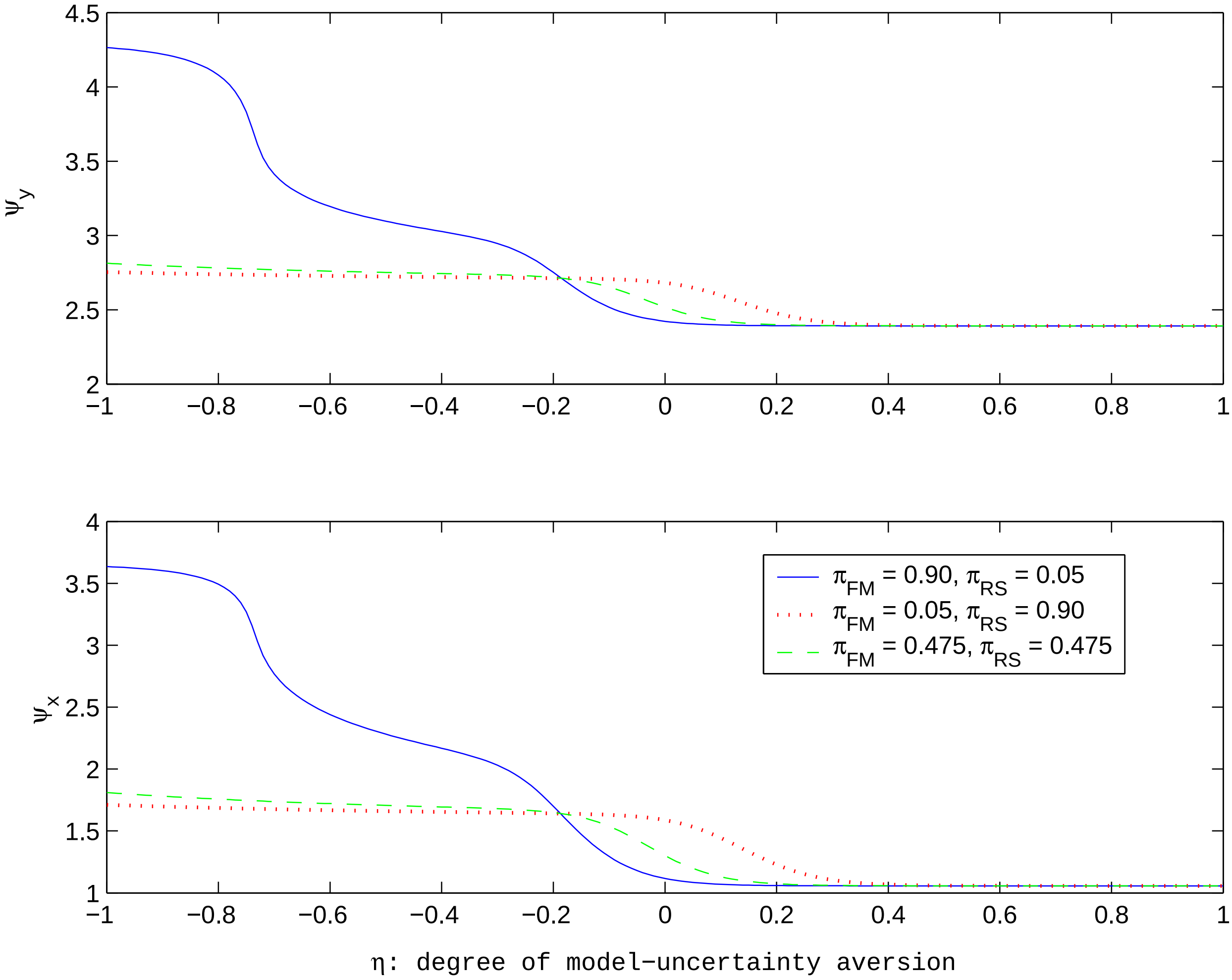


Figure 10: How the Taylor rule varies with the degree of model-uncertainty aversion when there is model-parameter uncertainty

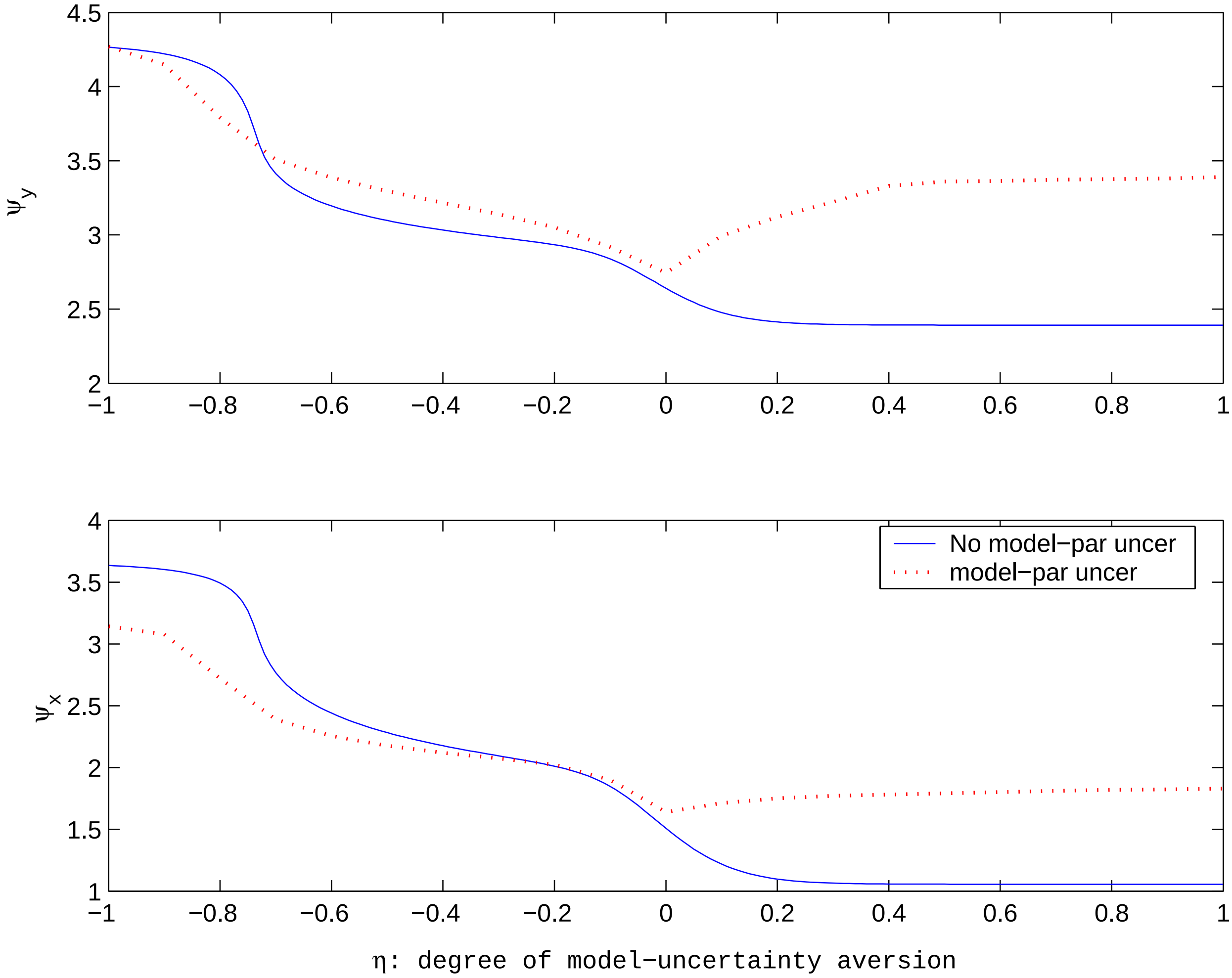


Figure 11: How the Taylor rule varies with the degree of aversion when the parameters of the measurement-error processes are known or not known precisely

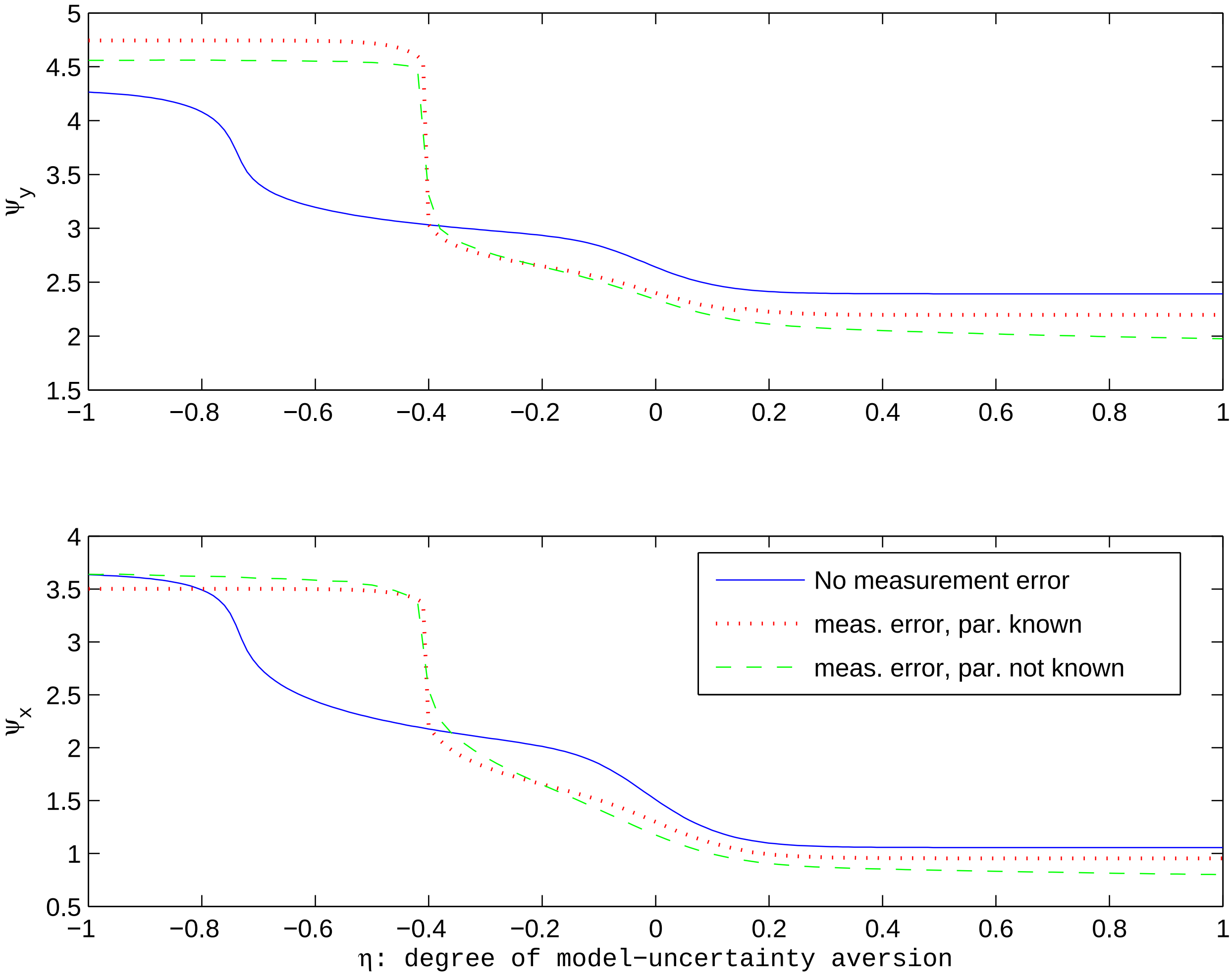


Figure 12: Real-time and final estimates of output gap. The real-time data are from the Federal Reserve Greenbook and the final data are based on 2001 data (HP filter).

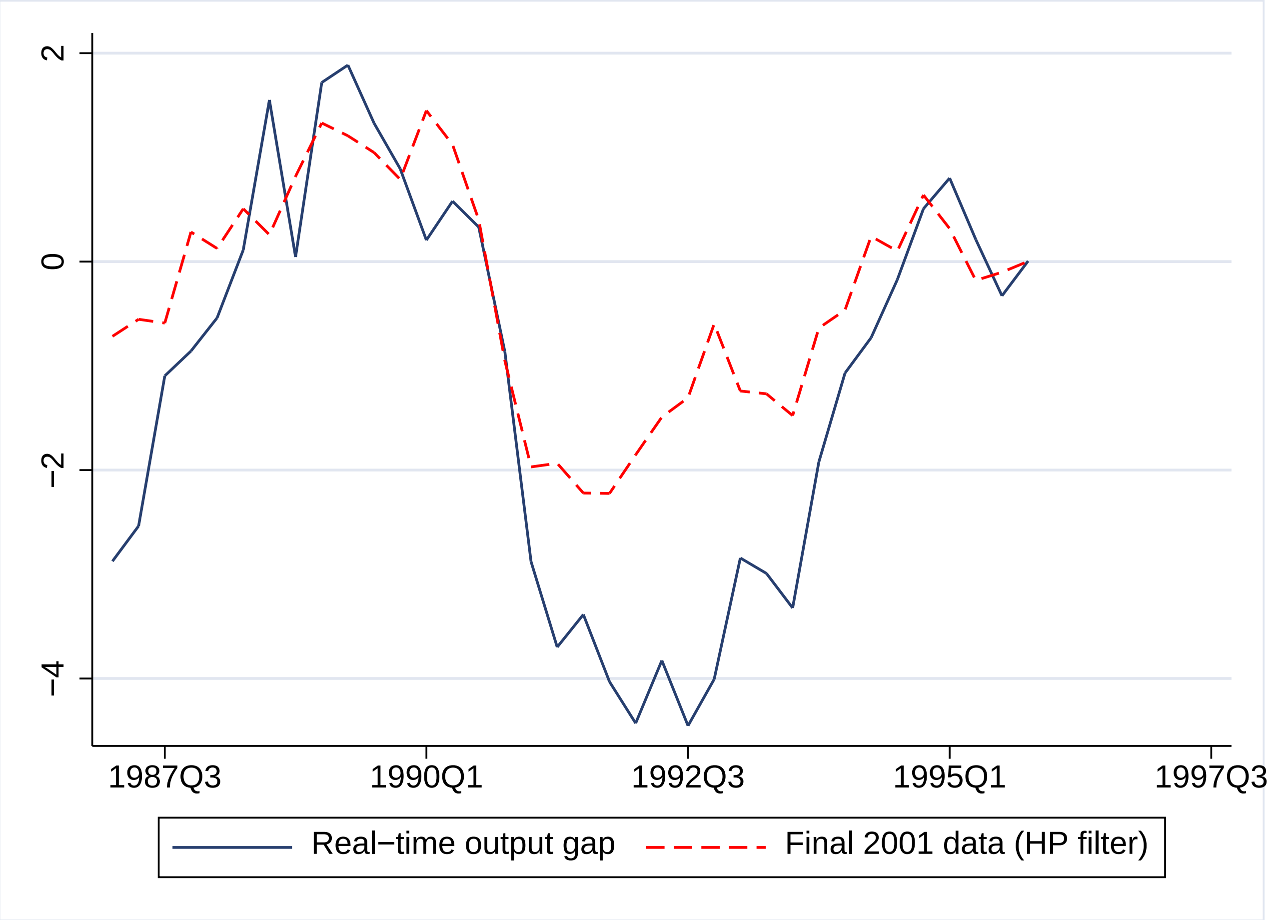


Figure 13: Real-time and final estimates of four-quarter average inflation. The real-time data are from the Federal Reserve Greenbook and the final data are based on 2001 data.

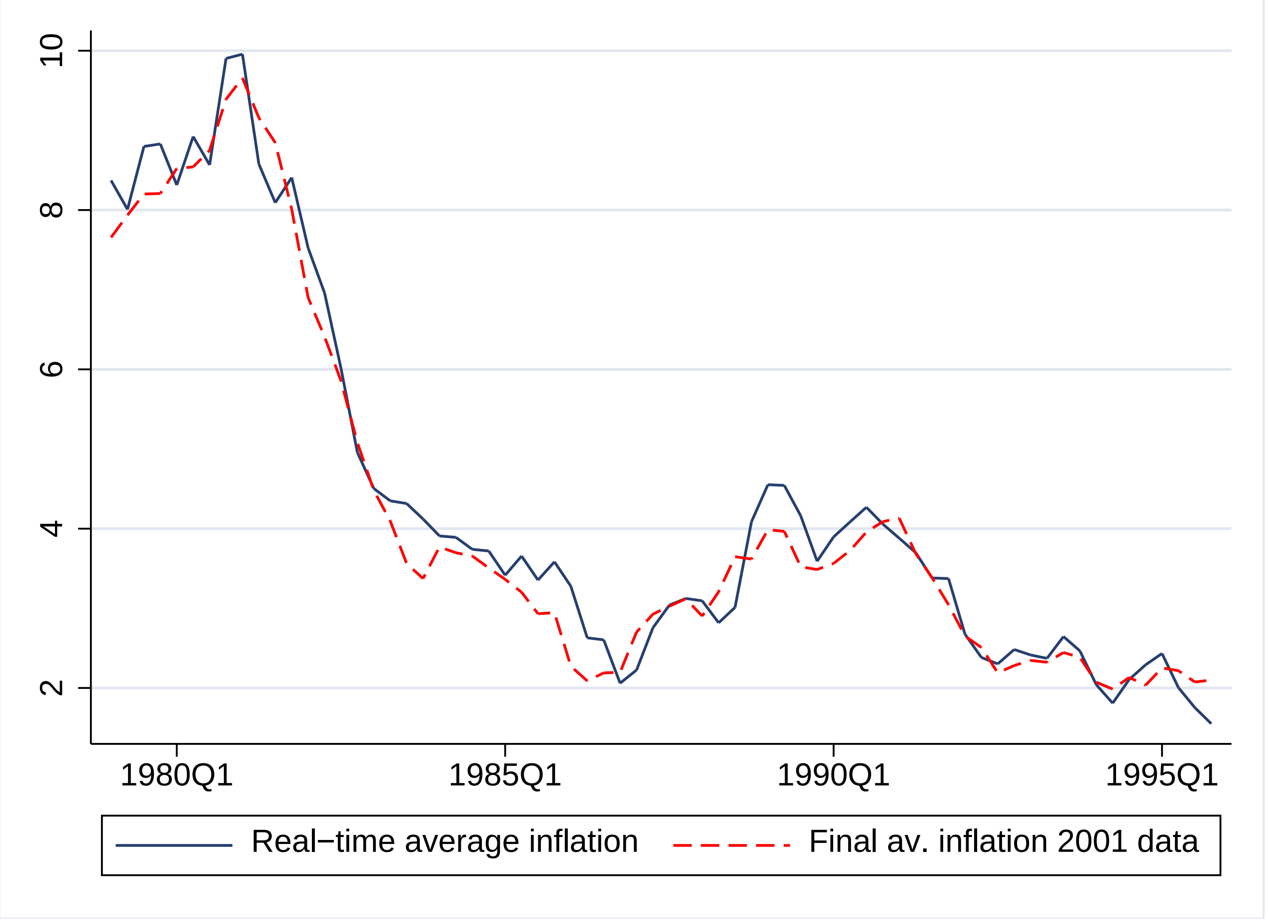


Figure 14: How the generalized Taylor rule changes with the degree of aversion

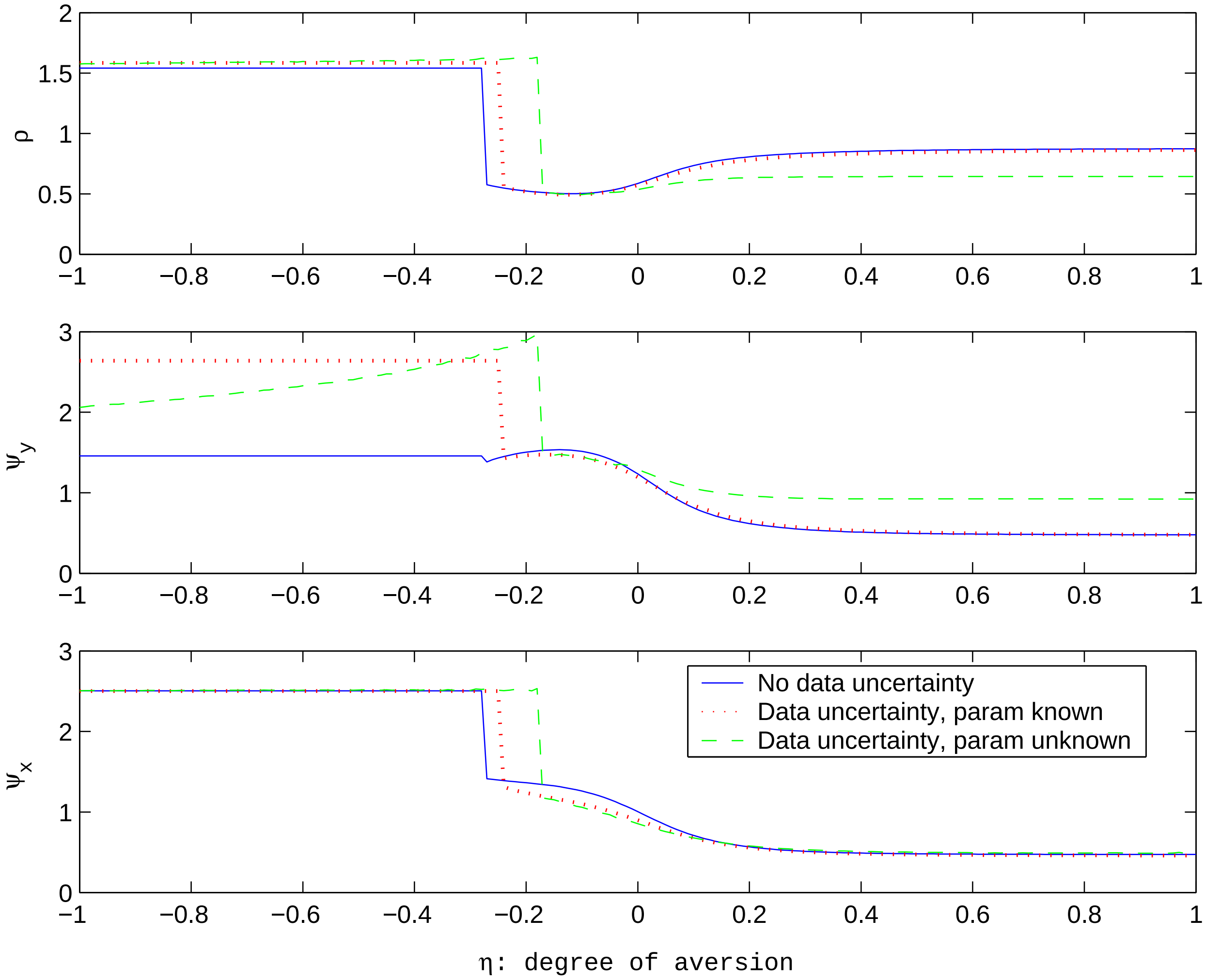
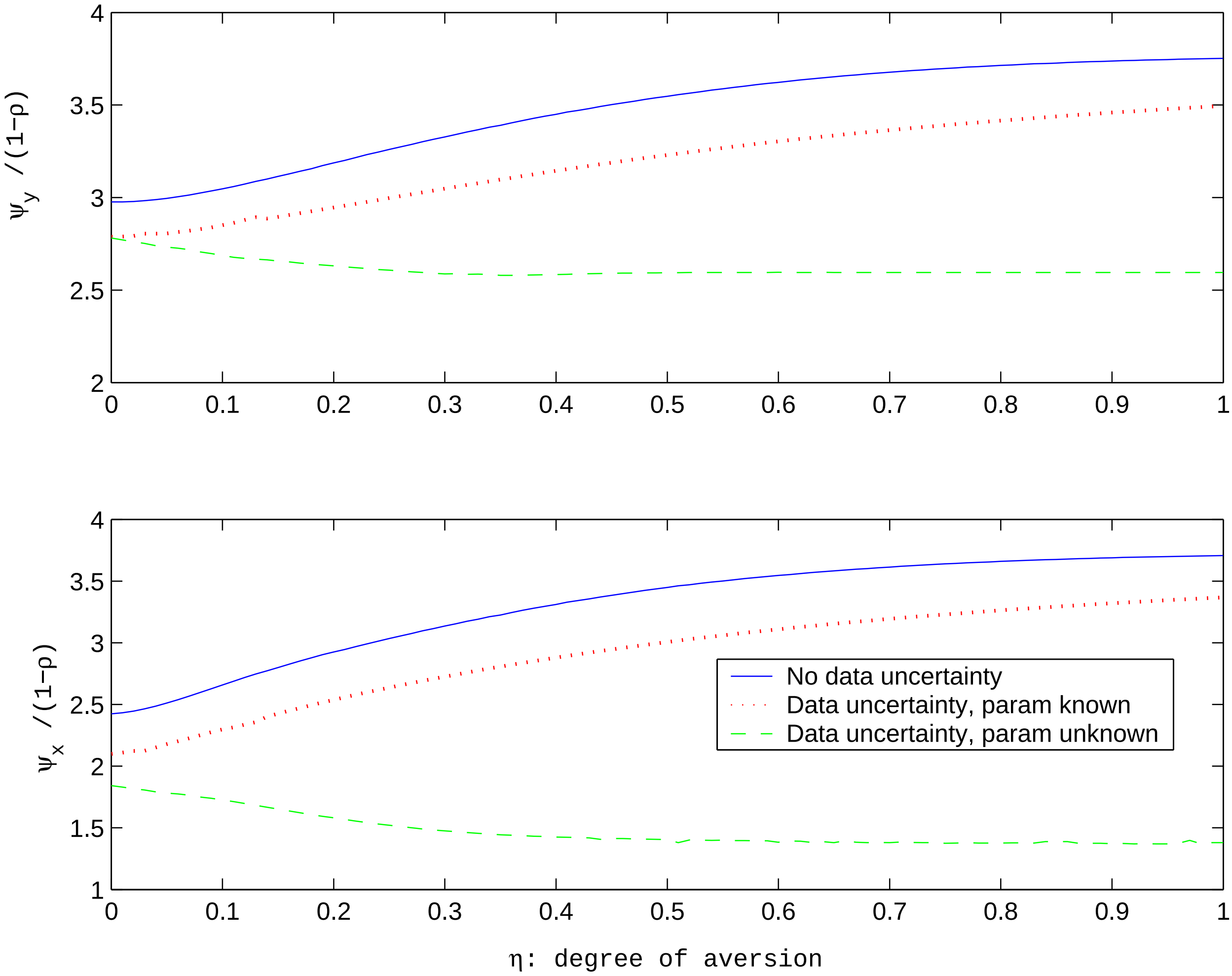


Figure 15: How the steady-state responses to inflation ($\frac{\psi_y}{1-\rho}$) and the output gap ($\frac{\psi_x}{1-\rho}$) change with the degree of aversion



Appendix A. Limits of Klibanoff, Marinacci, and Mukerji's Preferences

A.1 Limits under model uncertainty

Recall that my decision-making set-up when there is model uncertainty assumes that the agent chooses γ to minimize

$$h(\gamma) = \sum_{i=1}^{I} \pi_i \phi\left(v_i(\gamma)\right). \tag{35}$$

I claim in the main text that this is a general framework that nests as special cases the average-loss approach and the worst-case approach. That the average-loss approach is a special case of (35) is trivial; this follows by assuming a linear ϕ. In this appendix, I derive the worst-case approach as the special case when the degree of model-uncertainty aversion tends to infinity. I also emphasize the steps used to obtain the result. A more rigorous proof that adapts the arguments in Klibanoff, Marinacci, and Mukerji (2002) to my case can be similarly obtained.

Recall that, since I minimize loss, model-uncertainty aversion is characterized by *increasing and convex* ϕ. Suppose, then, that ϕ is increasing and convex to capture the attitude towards model uncertainty, and consider the function

$$\phi_n(x) = \phi(x)^n, \; n > 1.$$

For $x > 0$ (which is the relevant case, since losses are positive), $\phi_n(x)$ is an increasing and convex transformation of $\phi(x)$. Given

$$\frac{\phi_n''(x)}{\phi_n'(x)} = (n-1)\frac{\phi_n'(x)}{\phi_n(x)} + \frac{\phi''(x)}{\phi'(x)},$$

ϕ_n exhibits a greater degree of aversion to model uncertainty, which tends to infinity as n tends to infinity. Consider the objective function

$$\left(\sum_{i=1}^{I} \pi_i \phi_n\left(v_i(\gamma)\right)\right)^{\frac{1}{n}}. \tag{36}$$

(36) differs from (35) in that ϕ_n replaces ϕ and the objective is raised to the power of $\frac{1}{n}$. Since I am interested in choosing γ, the latter is just a monotonic transformation of the objective, and hence will not change the solution. By considering ϕ_n instead of ϕ, I can consider the effect of increasing model-uncertainty aversion by increasing n. Therefore, driving n to infinity:

$$\begin{aligned}
&\lim_{n\to\infty}\left(\sum_{i=1}^{I} \pi_i \phi\left(v_i(\gamma)\right)^n\right)^{\frac{1}{n}} \\
= &\lim_{n\to\infty} \phi\left(\max_i v_i(\gamma)\right)\left(\sum_{i=1}^{I} \pi_i \left(\frac{\phi(v_i(\gamma))}{\phi\left(\max_i v_i(\gamma)\right)}\right)^n\right)^{\frac{1}{n}} \\
= &\phi\left(\max_i v_i(\gamma)\right).
\end{aligned}$$

The second line follows because I assume that $v_i(\gamma)$ is bounded for each i. The third line follows since $\sum_{i=1}^{I} \pi_i \left(\frac{\phi(v_i(\gamma))}{\phi\left(\max_i v_i(\gamma)\right)}\right)^n$ tends to some finite number less than 1 and $\frac{1}{n}$ tends to 0 as $n \to \infty$.

Hence, by driving the degree of model-uncertainty aversion to ∞, I end up with a monotonic transformation of the worst-case objective function. Therefore, I can indeed conclude that Klibanoff, Marinacci, and Mukerji's (2002) preferences collapse to worst-case scenario preferences when the degree of model-uncertainty aversion tends to infinity.

A.2 Limits under model and parameter uncertainty

When an agent faces both model and parameter uncertainty, the criterion they minimize is

$$h^*(\gamma) = \sum_{i=1}^{I} \pi_i E_{\mathcal{P}_i} \; \phi\left(v_i(\theta_i, \gamma)\right). \tag{37}$$

As above, let

$$\phi_n(x) = \phi(x)^n, \; n > 1,$$

and consider

$$\left(\sum_{i=1}^{I} \pi_i E_{\mathcal{P}_i} \; \phi_n\left(v_i(\theta_i, \gamma)\right) \right)^{\frac{1}{n}}. \tag{38}$$

Therefore, driving n to ∞, I obtain

$$\begin{aligned}
& \lim_{n\to\infty} \left(\sum_{i=1}^{I} \pi_i E_{\mathcal{P}_i} \; \phi\left(v_i(\theta_i, \gamma)\right)^n \right)^{\frac{1}{n}} \\
= & \lim_{n\to\infty} \left(\sum_{i=1}^{I} \pi_i \int_{\theta_i \in \Theta_i} \phi\left(v_i(\theta_i, \gamma)\right)^n d\mathcal{P}_i \right)^{\frac{1}{n}} \\
= & \lim_{n\to\infty} \phi\left(\max_i \{ \max_{\theta_i \in \Theta_i} v_i(\theta_i, \gamma) \} \right) \\
& \cdot \left(\sum_{i=1}^{I} \pi_i \int_{\theta_i \in \Theta_i} \left(\frac{\phi(v_i(\theta_i, \gamma))}{\phi\left(\max_i \{ \max_{\theta_i \in \Theta_i} v_i(\theta_i, \gamma) \} \right)} \right)^n d\mathcal{P}_i \right)^{\frac{1}{n}} \\
= & \phi\left(\max_i \{ \max_{\theta_i \in \Theta_i} v_i(\theta_i, \gamma) \} \right).
\end{aligned} \tag{39}$$

Appendix B. Policy with a Generalized Taylor Rule

In the main text, I carry out my analysis with the Taylor rule as defined in Taylor (1993). Various authors have generalized the Taylor rule to allow it to depend on the lagged interest rate. In this appendix, I consider the effect of an aversion to model and data-parameter uncertainty when the central bank implements policy according to the generalized Taylor rule:

$$i_t = \rho i_{t-1} + \psi_y y_t + \psi_x x_t. \quad (40)$$

A casual OLS regression yields $i_t = 0.78 i_{t-1} + 0.41 y_t + 0.25 x_t$. Note that this implies, in the steady state, $i_t^* = 1.86 y_t^* + 1.14 x_t^*$.

Table 12 shows the optimal rule for a particular model and how it performs in the other models for the benchmark case $(\omega, \nu) = (1, 0.5)$. Note that, as for the Taylor rule without interest rate inertia, the generalized Taylor rule is less restrictive for the WG model. Indeed, the optimal rule leads to a loss of 7.7 in the WG model, whereas the RS and FM models' optimal rules lead to a loss of 13.81 and 24.73, respectively. The WG model, in fact, requires *superinertial* rules ($\rho = 1.56 > 1$). Superinertial rules, however, lead to instability in the RS model and a very high loss in the FM model. The RS model, in contrast, requires only a small degree of inertia ($\rho = 0.25$). Relative to optimum, it increases the loss in the WG model to 11.94 and the loss in the FM model to 28.18. Therefore, the optimal rule in the RS model does not do too badly in the other models. The FM model, on the other hand, requires a relatively high degree of inertia ($\rho = 0.87$). It increases the loss in the WG and RS models to 13.12 and 21.83, respectively.

So, how does model uncertainty affect the optimal generalized Taylor rule? Since the Taylor rule is less restrictive for the WG model, rules that work well in the WG model will be favoured at low degrees of aversion. In contrast, at higher degrees of aversion, rules that perform better in the FM model will be favoured. Figure 14 shows how the degree of inertia, response to inflation, and response to the output gap changes as I increase the degree of aversion. First, I will focus on the case where there is no data uncertainty (solid line). As expected, it takes a very "optimistic" central bank to choose superinertial rules. Indeed, only when the degree of aversion is negative does the central bank give more weight to the WG model. As the central bank becomes more averse, it gives up on superinertial rules in favour of rules with a degree of inertia less than one. Relative to the Bayesian framework ($\eta = 0$), I find that the central bank increases the degree of inertia as I increase the degree of aversion. Therefore, the central bank seeks robustness by making its policy instrument, the interest rate, more persistent. This is in line with the observation that there is reasonable interest rate smoothing in the data.

The dotted lines in Figure 14 refer to the case where there is data uncertainty and the measurement-error processes are known by the central bank. I find that measurement error does not change the profiles for the parameters of the generalized Taylor rule either qualitatively or quantitatively. The dashed lines, on the other hand, refer to the case where the parameters of the measurement-error processes are not known precisely. The most important difference in this case is that, relative to the Bayesian case, the central bank increases the degree of inertia more slowly. The implication of this slower increase in the degree of inertia is that the steady-state response to inflation ($\frac{\psi_y}{1-\rho}$) and the output gap ($\frac{\psi_x}{1-\rho}$) declines with the degree of aversion (see Figure 15). These steady-state values for the benchmark case are still more aggressive than observed in the data (2.59 vs 1.86 for inflation and 1.38 vs 1.14 for the output gap), but it is encouraging to note that, in a world of model and data-parameter uncertainty, increasing the degree of aversion to uncertainty reduces these steady-state values.

Appendix C. Solving the Three Models

C.1 Solving Rudebusch and Svensson (1999)

C.1.1 Transition Equation

Let $X_t = (y_t, y_{t-1}, y_{t-2}, y_{t-3}, x_t, x_{t-1}, i_{t-1}, i_{t-2}, i_{t-3})'$ be the state vector at time t and $e_{t+1} = (e_{\pi,t+1}, e_{x,t+1})'$ be the shock process. Then, given the Taylor rule $i_t = \psi_y y_t + \psi_x x_t$, the transition equation can be written as

$$X_{t+1} = AX_t + Ce_{t+1}, \tag{41}$$

where

$$A = \begin{pmatrix} b_0 & b_1 & b_2 & b_3 & b_x & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{d_i}{4}(1-\psi_y) & \frac{d_i}{4} & \frac{d_i}{4} & \frac{d_i}{4} & d_0 - \frac{d_i}{4}\psi_x & d_1 & -\frac{d_i}{4} & -\frac{d_i}{4} & -\frac{d_i}{4} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \psi_y & 0 & 0 & 0 & \psi_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

and $b_3 = 1 - b_0 - b_1 - b_2$.

Also, let

$$\Sigma_e = \begin{pmatrix} \sigma_\pi^2 & 0 \\ 0 & \sigma_x^2 \end{pmatrix}$$

be the covariance matrix of the shocks.

C.1.2 Loss function

First collect the variables appearing in the loss function in $l_t = (y_t, x_t, i_t) = DX_t$, where

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \psi_y & 0 & 0 & 0 & \psi_x & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Let

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \nu \end{pmatrix}$$

be the matrix of weights given to inflation, output-gap, and interest rate variability. Then, the loss function,

$$\text{var}(y_t) + \omega \text{var}(x_t) + \nu \ \text{var}(i_t),$$

can be written as

$$E(l_t K l_t{}') = \text{trace}(KD\Sigma_X D'),$$

where $\Sigma_X = E(X_t X_t{}')$ is the unconditional variance of the state vector that solves the Sylvester equation:

$$\Sigma_X = A\Sigma_X A' + C\Sigma_e C'.$$

C.2 Solving Woodford and Giannoni's model

C.2.1 Expectations in Woodford and Giannoni's model

To start, let $Y_t = (y_{t-1}, y_{t-2}, y_{t-3}, u_t, r_t^n, i_{t-1})'$ be the vector of predetermined state variables at time t. Let $Z_t = (y_t, x_t)$ be the vector of forward-looking variables at time t and $\epsilon_t = (\epsilon_{u,t}, \epsilon_{r,t})$. Given (7), (8), (9), (10), and (11), the WG model can be written compactly as

$$A_0 \begin{pmatrix} Y_{t+1} \\ E_t Z_{t+1} \end{pmatrix} = A_1 \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} + C_0 \epsilon_{t+1}. \tag{42}$$

Upon inversion of A_0 (A_0 is invertible for the chosen parameter ranges), I can write the above as

$$\begin{pmatrix} Y_{t+1} \\ E_t Z_{t+1} \end{pmatrix} = \begin{pmatrix} A_{YY} & A_{YZ} \\ A_{ZY} & A_{ZZ} \end{pmatrix} \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} + \begin{pmatrix} C \\ 0 \end{pmatrix} \epsilon_{t+1}. \tag{43}$$

In Woodford and Giannoni's model, inflation and output are endogenous forward-looking variables and have to be solved for in terms of the predetermined state variables. As in Currie and Levine (1993), solving (8) and (9) forward, I realize that inflation and output must be linear in the predetermined state variables:

$$Z_t = N_t Y_t. \tag{44}$$

From (44), it follows that expectations of inflation and output are given by

$$E_t Z_{t+1} = N_{t+1} E_t Y_{t+1}. \tag{45}$$

Taking expectations of (43), I obtain

$$E_t Y_{t+1} = A_{YY} Y_t + A_{YZ} Z_t, \tag{46}$$

$$E_t Z_{t+1} = A_{ZY} Y_t + A_{ZZ} Z_t. \tag{47}$$

Multiplying (46) throughout by N_{t+1}, equating the result with (47), and substituting out Z_t according to (44), I obtain:

$$N_t = (A_{ZZ} + N_{t+1} A_{YZ})^{-1} (N_{t+1} A_{YY} - A_{ZY}). \tag{48}$$

I then iterate on equation (49) until convergence, to obtain

$$N^* = N_t = N_{t+1}. \tag{49}$$

C.2.2 Transition equation for predetermined state variables

With the expectations of inflation and output given by (45) and (49), I substitute for the forward-looking variables, Z_t, in (43) to obtain:

$$Y_{t+1} = (A_{YY} + A_{YZ} N^*) Y_t + C \epsilon_{t+1}. \tag{50}$$

C.2.3 Loss function

Using the equations above, I solve for inflation and output to obtain $Z_t = (y_t, x_t)' = N^* Y_t$. Since $i_t = \psi_y \bar{y}_t + \psi_x x_t$, I can substitute out $(y_t, x_t)' = N^* Y_t$ to obtain i_t in terms of the predetermined variables only. Suppose this is $i_t = h Y_t$. Define

$$\begin{pmatrix} y_t \\ x_t \\ i_t \end{pmatrix} = D Y_t = \begin{pmatrix} N^* \\ h \end{pmatrix} Y_t,$$

and proceed in the same fashion as in section C.1.2. The unconditional variance of $(y_t, x_t, i_t)'$ similarly solves a Sylvester equation.

C.3 Solving Fuhrer and Moore's (1995) model

To solve Fuhrer and Moore's model, I adopt the methodology described in Anderson and Moore (1985), using the associated matlab code that can be found on the AIM Federal Reserve Board website (http://www.federalreserve.gov/pubs/oss/oss4/aimindex.html).

Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue*.**

2005

2005-5	Y a-t-il eu surinvestissement au Canada durant la seconde moitié des années 1990?	S. Martel
2005-4	State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?	P.J. Klenow and O. Kryvtsov
2005-3	Pre-Bid Run-Ups Ahead of Canadian Takeovers: How Big Is the Problem?	M.R. King and M. Padalko
2005-2	The Stochastic Discount Factor: Extending the Volatility Bound and a New Approach to Portfolio Selection with Higher-Order Moments	F. Chabi-Yo, R. Garcia, and E. Renault
2005-1	Self-Enforcing Labour Contracts and the Dynamics Puzzle	C. Calmès

2004

2004-49	Trade Credit and Credit Rationing in Canadian Firms	R. Cunningham
2004-48	An Empirical Analysis of the Canadian Term Structure of Zero-Coupon Interest Rates	D.J. Bolder, G. Johnson, and A. Metzler
2004-47	The Monetary Origins of Asymmetric Information in International Equity Markets	G.H. Bauer and C. Vega
2004-46	Une approche éclectique d'estimation du PIB potentiel pour le Royaume-Uni	C. St-Arnaud
2004-45	Modelling the Evolution of Credit Spreads in the United States	S.M. Turnbull and J. Yang
2004-44	The Transmission of World Shocks to Emerging-Market Countries: An Empirical Analysis	B. Desroches
2004-43	Real Return Bonds, Inflation Expectations, and the Break-Even Inflation Rate	I. Christensen, F. Dion, and C. Reid
2004-42	International Equity Flows and Returns: A Quantitative Equilibrium Approach	R. Albuquerque, G. Bauer, and M. Schneider

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :

Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
E-mail: publications@bankofcanada.ca
Web site: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca