Macroeconomic Consequences of Optimal Information Acquisition

Ian Corrick

PhD

University of York

Department of Economics and Related Studies

September 2013

Abstract

This thesis considers the extension of the 'sticky information' concept of Mankiw and Reis (2002), by which agents form expectations rationally conditional on out-of-date information, with new information arriving probabilistically, to models where the probability of receiving new information is a choice variable. Previously this has been done only for simple, restricted economic models (e.g. Branch et al. (2009)), but not for DSGE models as used in modern macroeconomic theory.

Numerical results for two different models are presented and then estimation of the more fully-featured model is conducted.

It is found that for a simple model of monopolistically competitive firms the introduction of endogenous sticky information can lead to multiple equilibria, particularly when there is strong strategic complementarity. The optimal updating probabilities are strongly responsive to the variability of monetary policy shocks and the macroeconomic effect of this is that changes in shock variability cause a trade-off between the variance of output and inflation, which doesn't occur in the case of exogenous sticky information.

For a DSGE model with endogenous sticky information with more than one type of agent, with agent types having separate updating probabilities, it is found that the possibility of multiple equilibria is reduced relative to the simple model. The numerical results for this model show that the effects of changes in the coefficients of monetary policy on the volatility of variables and on the model dynamics, as characterized by the impulse response functions, may be amplified by the presence of endogenous sticky information.

Finally, the estimation results show there has been a reduction in the updating probability of firms, and an increase in that of households, in the period 1984Q3 - 2006Q1 compared to 1955Q1 - 1984Q2. This can be explained by a switch of monetary policy responsiveness from the output gap to inflation and by changes in the volatilities of the exogenous shocks themselves.

Contents

List of Figures

List of Tables

Acknowledgements

My supervisors, Professor Peter Smith and Professor Neil Rankin, have provided much helpful advice, guidance and encouragement throughout my studies, for which I am grateful.

I would also like to thank Professor Subir Chattopadhyay for being a member of my Thesis Advisory Panel.

Thanks are also due to the other members of the doctoral student cohort in the Department of Economics, who have contributed to making my time at York enjoyable.

Author's Declaration

All of the material contained herein is original and single-authored. No parts of this thesis have been submitted for a degree at any other institution.

Chapter 1

Introduction

Since the introduction of rational expectations to macroeconomic theory, there has been a great deal of research into whether the requirements of this assumption are met and if not how the specification of the constraints upon agents' decision-making can be framed in order to provide a more accurate and realistic description. In addition the effects of such constraints upon general economic behaviour have been extensively studied. This area of research has been especially actively pursued in the last fifteen or so years, with many alternative means of modelling the limited capacity of economic agents to process information being suggested. These alternatives include signal extraction under a limited capacity information channel, known as 'rational inattention' and introduced by Sims (2003) and least squares/adaptive learning under which agents behave as econometricians, updating their beliefs about the economy as more data arrives (this topic is covered, for example, by Evans and Honkapohja (2009)). Both of these approaches have had some success in capturing empirical behaviour and are to some extent plausible. However, the type of information constraint which this thesis considers is the 'sticky information' concept introduced by Mankiw and Reis (2002). This assumes that all agents form their expectations rationally, but that these expectations are conditional on information sets which are not up to date, so that while agents do act upon the mathematical expectation of future variables, they may be unaware of recent realizations of the variables which other agents do know about. This raises the possibility of strategic interactions between agents and this possibility is considered here.

A general goal of constrained information modelling, in addition to the microeconomic objective of accurately describing the way in which economic decisions are made, is to find a mechanism by which macroeconomic models produce dynamics which accord with the data. In particular, the key finding that the responses of macro-variables to exogenous shocks are gradual and persistent rather than instantaneous has led to many modifications of classical business cycle models, which cannot replicate this fact. The ability of models in which agents are informationally constrained to produce persistent dynamics is one of the reasons for the continuing interest in them. This important property will be considered repeatedly in the chapters which follow.

The contribution of this thesis is to consider the characteristics of models in which decision-making is subject to sticky information and the probability of receiving new information (the 'stickiness' parameter) is determined endogenously as an optimal choice by agents. To date this has only been done in either very simple models or, in one case (Drager (2011)), in a DSGE model but with several additional restrictions required for tractability, such that the generality of the results may be limited. Here the sticky information mechanism is endogenized with minimal restrictions and therefore the results obtained are all original.

The following chapters provide a number of new results concerning the endogenous sticky information model. Chapter 2 analyses a simple model of monopolistically competitive firms with exogenous aggregate demand, where it is found that for a plausible degree of strategic complementarity in price-setting the model may have multiple equilibria and, in particular, there are multiple equilibria of the updating probabilities which are consistent with the other calibrated structural parameters. It is also found that the choice of updating probability depends quite strongly on the cost of receiving and processing new information, and on the variability and persistence of the exogenous shocks, with the relationships being those which would be expected intuitively. When changes to the volatility of the exogenous monetary shock are considered, it is found that there may be a trade-off between reducing the variance of inflation and increasing the variance of output, and vice versa. This trade-off does not occur in the model with exogenous sticky information, as multiple equilibria do not exist in that model. This is a significant finding, in that it demonstrates that the possibility that agents may respond to policy changes by changing the average frequency with which they update their plans has further implications for the conduct of policy, beyond those which are apparent in the exogenous sticky information model. However, this finding may be limited by the relative simplicity of the model, which does not contain several of the features which are common in the type of fully-specified macroeconomic model used for monetary policy analysis.

Chapter 3 constructs and analyses a model which has greater similarity to modern DSGE models (and indeed, is an example of such). The particular model followed is that of Mankiw and Reis (2006b), with the additional feature that the updating probabilities are endogenously determined. The model features two different types of informationally-constrained agent, firms and households, each with a separate updating probability. Several aspects of the model are common to the New Keynesian literature, for example price-setting firms are monopolistically competitive and these hire labour from wage-setting trade unions/households. Monetary policy can be represented by a Taylor rule which responds to inflation and the output gap but does not depend upon lagged interest rates. Additional exogenous shocks to the model enter through aggregate demand/fiscal policy, the labour substitution elasticity and the consumption substitution elasticity. The chapter finds that the optimal updating probabilities of different agent types display an asymmetric strategic inter-relationship, with the choice of firms depending quite strongly on the choice of households but relatively weak dependence of the household choice upon the probability chosen by firms. The model does not appear to possess multiple equilibria, at least under the range of parameterizations investigated, which may be due to the existence of these inter-agent-type relationships. It is also found that the optimal updating probabilities vary with the costs of information, the standard deviations of the exogenous shocks and the coefficients of the monetary policy rule.

Section 3.5 of the third chapter shows that numerical experiments with this model reveal that when there is an increase in the weight attached to inflation by monetary policy (with an assumed target of zero), the effect varies between the exogenous and endogenous sticky information models. For exogenous sticky information it is found that an increase in ϕ_p (the inflation weight in monetary policy) reduces the volatilities of both output and inflation. However, when the same change is considered under endogenous sticky information it is found that the effects of the policy change are amplified with respect to inflation, as the change in inflation volatility is greater than in the exogenous sticky information case; with respect to output the effect of the policy change is reduced, with a smaller decrease in output variability. For changes in the volatility of monetary policy shocks, σ_ε, it is found that again the presence of endogenous sticky information has a significant effect. As the volatility increases the updating probabilities of both types of agents increase, and in the case of firms the optimal choice is eventually to be always fully-informed. In the case of exogenous sticky information, when σ_ε increases the variances of both output and inflation increase, and when endogenous sticky information is allowed for both of these variances increase by a larger amount, as the increase in updating probabilities means that prices, wages and consumption respond much more quickly to the effects of shocks. Hence any change in monetary policy which corresponds to a reduction of the accuracy of the Taylor rule as a representation of the implementation of monetary policy will be amplified with respect to the variability of output and inflation. Concerning the dynamics of the economy, the results of the numerical simulations of the impulse response functions of output and inflation in chapter 3 reveal two aspects of the effects of monetary policy changes. Firstly, an increase in the weight attached to inflation in monetary policy results in agents updating less frequently on average, with the effect that both inflation and output respond less strongly to shocks, with smaller maximal responses and generally smaller responses for several quarters after the impact date. A second result is that a change in endogenous updating probabilities, induced by a change in monetary policy coefficients, could change the length of time before the maximal response occurs, i.e. the impulse response functions peak at different dates. While this effect is not very strong in terms of the shift of the peak response, combined with the change in the size of the responses this is seen to lead to quite different dynamics compared to the benchmark case.

Chapter 3 also considers the dynamics of output and inflation in response to a monetary policy shock when the standard deviation of the shock changes. In the exogenous sticky information case such a change would merely rescale the impulse response function. However, it is found that in the endogenous sticky information case both the size of the response (in addition to the rescaling) and the shape of the response change. For a decrease in σ_ε both types of agent choose to update less frequently and consequently the responses of output and inflation to shocks are smaller and, in the case of output, the peak of the response occurs later than under the benchmark parameterization. For an increase in σ_ε both types of agent choose to update more frequently and consequently the immediate responses of output and inflation to a shock are much greater and the peaks of the response functions occur sooner.

Overall, the numerical results in chapter 3 suggest that the addition of endogenous sticky information to a model is highly significant, with the properties of output and inflation being very different as a result. To the extent that this specification of the informational limitations faced by agents when forming their expectations is accurate, the results suggest that the addition of this feature to macroeconomic models is worthwhile, as its omission does change the properties of the model.

Having analysed the theoretical behaviour of a DSGE model featuring endogenous sticky information, the next stage in the research was to attempt to estimate this model in order to evaluate whether it does provide an improvement in empirical performance. The empirical methods used are the extended method of simulated moments (which may now also be more commonly known as indirect estimation) and structural vector auto-regressions, where the identification of shocks is based on the signs of the derived impulse responses. The results from the indirect estimation, presented in chapter 4, show that there have been some changes in the updating behaviour of private agents within the USA economy over the period 1955 - 2006 and suggest some of the possible reasons for this, by estimating the model over two sub-samples, with 1984 as the split date. After 1984, firms have chosen much lower updating probabilities while households have chosen much higher updating probabilities. This finding is quite surprising, but the results also indicate why this may have happened. The first potential explanation is that monetary policy became more responsive to inflation and less responsive to the output gap, which may have reduced the exposure of firms to variability of prices but increased the exposure of households to variability of consumption. A second possible explanation is that the estimation results show a slight decrease in the standard deviation of shocks to the goods substitution elasticity, σ_ν, and a large increase in the standard deviation of shocks to the labour substitution elasticity, σ_γ. These may have lead to less volatile price mark-ups but more volatile wage mark-ups, allowing firms to update less frequently without experiencing large expectational errors, but requiring households to update more frequently in order to reduce the increased expectational errors they might face. These two explanations are complementary and accord with the theoretical results found in chapter 3. The sensitivity of the results to the particular method of data filtering/detrending used is also considered in chapter 4, where it is shown that the results are somewhat sensitive, although the logic of the updating behaviour is still satisfied.

Chapter 4 also considers the ability of the estimated model to match the dynamic responses of variables to exogenous shocks. This is done using the SVAR method. These results do not suggest that the estimated model performs entirely satisfactorily empirically, as few of the empirical impulse responses are closely matched by the estimated model responses. Some of the possible reasons for this which are considered are that the identification method may have failed to fully identify the empirical counterparts of the structural shocks and alternatively that the comparison may have been somewhat biased towards 'rejection' because it was not possible to produce confidence bands for the impulse responses. These possibilities and possible remedies to these are discussed in the concluding section of chapter 4.

Finally, chapter 5 provides an overall conclusion and summary of the research findings and also discusses possible future extensions of this work or other related possibilities.

Chapter 2

Endogenous Information Updating In The Mankiw-Reis Framework

2.1 Introduction

The role of incomplete information in the decision-making of agents has recently attracted much attention in macroeconomics, with applications, for example, to price-setting in response to shocks (Mackowiak and Wiederholt (2009)), the effectiveness of monetary policy (Cone (2008)) and price-setting in heterogeneous agent models (Coibion and Gorodnichenko (2011)). A major stimulus to much of this work was Mankiw and Reis (2002), which introduced the idea of sticky information as an expectational analogue to the Calvo (1983) model of sticky prices[1]. Specifically, agents were modelled as only receiving new information infrequently, rather than always having up-to-date information, and setting prices optimally conditional upon the information they had. As specified by Mankiw and Reis (2002), the probability of receiving new information is exogenous and fixed. While this is sufficient for the purpose of generating impulse response functions typical of such a model and comparing these to those typical of Calvo sticky price behaviour, and demonstrating that the sticky information model can qualitatively replicate empirical relationships such as the "acceleration phenomenon" (ibid. p1311), it

[1]The other main approaches to incomplete-information modelling in macroeconomics are rational inattention (following Sims (2003)) and learning (see for example Evans et al. (2001)).

may be considered unsatisfactory that a key determinant of the outcome of the model is exogenous. Furthermore, endogenizing the probability of information updating may allow the investigation of the relationships between this and other model parameters and thereby permit further conclusions to be made about behaviour under the sticky information assumption.

Within the existing literature an endogenization of the information acquisition probability has been undertaken by Branch et al. (2009). The model used therein follows the model of sticky information under optimal monetary policy in Ball et al. (2005), but with the addition of endogenous determination of the distribution of information. The main consequence of this is that for certain ranges of policy-maker preferences on the volatilities of output and inflation the conventional trade-off between the two may not arise, i.e. increased weight on inflation volatility may lead to reduced volatility of both inflation and output, due to the effect of such policy action on the average rate at which private agents update their information. However, the Ball et al. (2005) optimal monetary policy is something of a special case, in which the policy-maker is able to credibly commit to a money/demand target one period in advance and thereby determine expectations of the one-period ahead price level. Arguably this is unsatisfactory, as it may be that under sticky information the possibility for effective policy surprises is greater, at least if the probability of agents receiving new information in each period is relatively small, so credible policy would be even more difficult than under full information . Here the model uses the basic framework of Ball et al. (2005) , but without the addition of a policy-maker, which avoids these issues of policy commitment and credibility. Although it would be desirable to include a realistic representation of the conduct of monetary policy (e.g. following a Taylor rule), this is not compatible with the model used below, which considers a static choice of probability. As in Branch et al. (2009) the model here requires sticky information to be characterized by a single probability applicable to all agents - hence it cannot be used to consider different sub-groups of agents re-optimizing over time subject to different information sets. This limitation renders the use in Branch et al. (2009)of comparative static results, which are derived from minimization of an asymptotic per-period loss function, to explain dynamic changes in the USA economy as somewhat doubtful, although it is possible that these results may be indicative of the eventual outcomes that would arise under a dynamic model.

The motivation for what follows is therefore both to consider whether the properties of the endogenous updating probability in the most general model are the same as those found under the restrictive conditions in Branch et al. (2009) and to determine whether in this general case there may be some parameter values for which there is no (comparative) trade-off between inflation and output volatility. This is also motivated by the question of whether the macroeconomic variables exhibit different behaviour under endogenous sticky information compared to the exogenous case, as considered originally by Mankiw and Reis (2002) and subsequently in applications of this concept, such as those noted above.

It will be shown that the qualitative relationships between the updating probability and the exogenous parameters seem to be the same in the general model as in Branch et al. (2009), but that the comparative statics results do not include any instances of a parameter variation resulting in the output and inflation variances increasing or decreasing together. A complication in these results is that it is found that in many cases, including those which could be argued to involve the most plausible parameter values, there may be multiple equilibria in the information acquisition probability, with the consequence that the comparative statics results are uncertain[2].

2.2 Model

Following Ball et al. (2005) the agents constitute a unit-length continuum of monopolistically-competitive 'yeoman farmers' indexed by i. The utility of agent i at time t is represented by

$$U_{it} = \frac{C_{it}^{1-\sigma}}{1-\sigma} - \frac{L_{it}^{1+\psi}}{1+\psi} \tag{2.1}$$

where L_{it} is the labour effort expended in the production of variety i consumption good and C_{it} is a constant elasticity of substitution aggregate of the goods produced by

[2]Branch et al. (2009) acknowledges the possibility of multiple equilibria but does not consider it in detail.

all other agents:

$$C_{it} = \left[\int_0^1 \left(C_{it}^j \right)^{\frac{\gamma-1}{\gamma}} dj \right]^{\frac{\gamma}{\gamma-1}} \quad (2.2)$$

where C_{it}^j is the consumption of good type j by agent i at time t. Individual production is simply given by[3]:

$$Y_{it} = AL_{it} \quad (2.3)$$

The optimization of the objective in equation 2.1 subject to the constraints of equation 2.3, a standard budget constraint and a cash-in-advance constraint is described in appendix A.1.1; this solution provides the optimal (log) price-setting rule for agents (under flexible prices)[4]:

$$p_{it}^* = p_t + \alpha y_t + u_t \quad (2.4)$$

u_t is assumed to follow the process $u_t = \rho u_{t-1} + \varepsilon_t$ with $\rho \in (0, 1)$ and $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$.

Price-setting is subject to the sticky information assumption, which is encapsulated by:

$$p_{it}^k = E_{t-k} \left[p_{it}^* \right], k = 0, 1, 2, ..., n \quad (2.5)$$

which is that a firm that last received new information k periods ago will set a price at time t equal to the rational expectation of the current optimal price, conditional on the time t - k information set. It could be argued that behaviour under the sticky information model fails to satisfy rationality. At the aggregate level it is certainly true that the overall outcome will differ from that which would occur under full rationality, and indeed one of the justifications for the model is that it allows the very strict requirement of total rationality to be relaxed in a manner which still allows individual decisions to be conditionally rational and allows the behaviour of the agents to be represented in a way which does not require any other significant changes to the model structure (an

[3]The technology parameter A is assumed to be constant over time and common across agents.

[4]The additional term u represents a mark-up shock. This ad-hoc addition (for the sake of congruency with other work) could be justified as resulting from variable taxes or stochastic preferences, although these are not modelled here.

advantage of which being that comparisons of models with full rationality and sticky information may be relatively straightforward). In this model even those agents who do receive full/current information in a period will not behave in exactly the same way as agents in the model of pervasive full information because they will allow for the decisions of those other agents who do not have the most recent information. By similar reasoning, it could be argued that knowledge of this interdependence on the part of agents who do not have full information will allow them to predict to some extent the behaviour of those agents who do have full information, which could be thought of as a further departure from full rationality. Again, such a departure could be justified on the grounds that it remains the case that individual agents are acting rationally, conditional on the available information, and provided that this is satisfied any departure from 'aggregate' rationality is to be welcomed in that it may allow the model to better fit those empirical regularities which the standard New Keynesian model is not able to. The empirical performance of the endogenous sticky information is considered in chapter 4. The receipt of new information is determined probabilistically, and assuming that this probability, λ, is equal across all agents results in the following equation for the (approximate) aggregate price index:

$$p_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \, E_{t-j} \left[p_t^* \right] \tag{2.6}$$

Aggregate demand is assumed to have the representation[5]:

$$y_t = m_t - p_t \tag{2.7}$$

where money/nominal demand is assumed to follow

$$m_t = \mu + \delta m_{t-1} + e_t \tag{2.8}$$

with $\delta \in (0,1)$ and $e_t \sim i.i.d.\ (0, \sigma_e^2)$.

In order to model the choice of an optimal λ it is assumed that agents choose the updating probability in order to minimize the discounted sum of the expected squared

[5]All variables in natural logs

deviations of the actual price set in each particular period from the optimal price (i.e. that which would be set if new information was received in every period) and that there is some cost to acquiring and processing information, such that these losses are not necessarily minimized by choosing $\lambda_i = 1$. As a necessary starting point for the model it is assumed that all agents make this one-off decision at some initial time t, at which they have full information. As a consequence of this the infinite history of the Mankiw and Reis (2002) model is truncated in the sense that all future prices will be based on information from time t or later, such that the aggregate price at time t + n for some n $\geq$ 1 is now given by the distribution of information from time t onwards[6]:

$$p_{t+n} = \lambda p^*_{t+n} + \lambda \sum_{j=1}^{n-1} (1-\lambda)^j E_{t-j}\left[p^*_{t+n}\right] + (1-\lambda)^n E_t\left[p^*_{t+n}\right] \tag{2.9}$$

Apart from this modification all of the other equations above continue to apply.

A distinction is made in the loss function between the individual updating probability λ_i (the decision variable for firm i) and the 'aggregate' $\overline{\lambda}$ which represents the choice of all other agents. The choice can be considered as a game between agents, as the loss to each agent depends on aggregate price, which in turn depends on $\overline{\lambda}$, although each agent takes $\overline{\lambda}$ to be invariant to their individual choice, given the monopolistic competition setting of the model. Given that agents are identical at time t and that as specified the model requires a single updating probability it is appropriate to focus on symmetric Nash equilibria of the game, i.e. those outcomes where $\lambda_i^* = \overline{\lambda}\ \forall i$, which allows the choices of other agents to be represented as a single aggregate value, $\overline{\lambda}$ [7]. In principle agents will consider the infinite future when making this decision, although for the computational results below it is necessary to set a finite horizon of N periods.

[6]The index i on the individual optimal price is ignored hereafter as agents are ex-ante identical

[7]The consideration of the problem as a game and the focus on symmetric Nash equilibria follows Branch et al. (2009).

Therefore the loss function for the agent, for some planning horizon N $\leq +\infty$, is

$$Loss_i = \beta\left(1-\lambda_i\right)E\left[\left(E_t\left[p^*_{t+1}\right]-p^*_{t+1}\right)^2\right] \quad + \textstyle\sum_{n=2}^{N}\beta^n\Bigg(\ \lambda_i\sum_{z=1}^{n-1}(1-\lambda_i)^{n-z}$$

$$\cdot E\left[\left(E_{t+z}\left[p^*_{t+n}\right]-p^*_{t+n}\right)^2\right]+\ \ (1-\lambda_i)^n E\left[\left(E_t\left[p^*_{t+n}\right]-p^*_{t+n}\right)^2\right]\Bigg)$$

$$+\sum_{n=1}^{N}\beta^n C\lambda_i^2 \qquad (2.10)$$

where $\beta \in (0,1)$ is the discount factor and C is the cost parameter. Here, as in Branch et al. (2009), C does not have a direct financial interpretation, but is rather some indirect measure of the per period costs associated with using new information, such as labour/time expended on gathering and processing new information[8], expressed in units compatible with the other terms. It is assumed that this cost parameter is constant across time and that the total per-period information cost is the product of C and the square of the updating probability, so the final term in equation 2.10 is the present value of the total information costs. The cost is specified as depending on the square of the probability on the grounds that some information is readily available but it may be increasingly costly to achieve 'full information' if this requires an increasing proportion of managerial labour or, from the perspective of the household, if this requires extra effort which reduces utility in a similar manner to labour effort (although this is not explicitly modelled). A possible mathematical consequence of the total cost depending quadratically on the probability is that the solution may be more likely to be interior, as at low probabilities the total cost may be small relative to the benefits of lower weighted MSE, whereas at probabilities close to one the total cost will be relatively much larger and potentially large relative to the benefits of lower weighted MSE; overall this may make corner solutions relatively unlikely, although it will not rule them out entirely. The first term in equation 2.10 is the expected loss in period t + 1. With probability λ_i the agent will receive new information and set the optimal price, incurring no loss. With probability (1 - λ_i) the agent will not receive new information and so will set the price

[8]In a more realistic model these costs could for instance be fees paid to external analysts/forecasters, or some proportion of the wages paid to managerial employees.

equal to the conditional expectation of the optimal price given time t information. Hence the expected squared deviation of this actual price from the optimal price, multiplied by the probability of setting this actual price and the discount factor, is the appropriate loss value. The second term in equation 2.10 contains the equivalent values for periods t + 2 to t + N, where the first term in the bracketed expression under this sum is the sum of the losses arising from setting actual price at date t + n based on information received at date t + z, multiplied by the probability of setting such a price and the appropriate discount factor. The last term in this bracketed expression is the probability-weighted loss arising in period t + n if no new information is received after period t.

Using equations 2.4, 2.7, 2.8 and 2.9 it is possible to solve for the expectations terms (see appendix A.1.2) to arrive at the following representation of the loss function

$$
\begin{aligned}
Loss_i = \sum_{n=1}^{N} \beta^n C \lambda_i^2 + \beta\left(1-\lambda_i\right)\left(\frac{\alpha^2\sigma_e^2+\sigma_\varepsilon^2}{(1-(1-\alpha)\overline{\lambda})^2}\right) + \\
\sum_{n=2}^{N} \beta^n \Bigg(\lambda_i \sum_{z=1}^{n-1}(1-\lambda_i)^{n-z} \sum_{b=0}^{n-z-1}\Bigg(\Big(\frac{-\alpha\delta^b}{1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j}\Big)^2 \sigma_e^2 \\
+\Big(\frac{-\rho^b}{1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\lambda)^j}\Big)^2 \sigma_\varepsilon^2\Bigg) \\
+(1-\lambda_i)^n \sum_{b=0}^{n-1}\Bigg(\Big(\frac{-\alpha\delta^b}{1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j}\Big)^2 \sigma_e^2 \\
+\Big(\frac{-\rho^b}{1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j}\Big)^2 \sigma_\varepsilon^2\Bigg)\Bigg)
\end{aligned}
\tag{2.11}
$$

From this it is possible to determine the effect of some of the underlying parameters on the agents' expected loss, although not necessarily the effect of these on the equilibrium updating probability (in the next section the one-period problem is considered, where it is possible to solve analytically for the optimal individual λ). It can be seen above that the loss is increasing in the persistence and volatility of the fundamental shocks (δ, ρ and $\sigma_e^2, \sigma_\varepsilon^2$ respectively). This accords with the intuition that knowledge of shock realizations is expected to be more valuable the more prolonged are their effects and the greater they are likely to be.

An alternative formulation of the loss, if the agent was choosing λ in order to minimize the expected per-period squared price deviations subject to the cost, given an indefinite decision date, (which would be analogous to that of Branch et al. (2009)), would be:

$$Loss_i = \lambda_i \sum_{k=0}^{\infty} (1-\lambda_i)^k E\left[(E_{t-k}[p_t^*] - p_t^*)^2 \right] + C\lambda_i^2 \tag{2.12}$$

The meaning of the terms in this is the same as for equation 2.10, and the expectational error terms could again be replaced with the appropriately - dated equivalents to the terms in equation 2.11.

2.3 Analysis of One-Period Problem

In order to facilitate an analytical determination of the relationships between the endgoenous updating probability and the other parameters of the loss function, and the determination of the existence and multiplicity of equilibria, consider the loss function for an agent looking merely one period ahead:

$$Loss_i = \beta (1-\lambda_i) \left(\frac{\alpha^2 \sigma_e^2 + \sigma_\varepsilon^2}{(1-(1-\alpha)\overline{\lambda})^2} \right) + C\lambda_i^2 \tag{2.13}$$

Minimizing this loss with respect to the individual updating probability:

$$\begin{aligned} \frac{\partial Loss_i}{\partial \lambda_i} &= -\beta \tfrac{\alpha^2 \sigma_e^2 + \sigma_\varepsilon^2}{(1-(1-\alpha)\overline{\lambda})^2} + 2C\lambda_i = 0 \\ \Rightarrow \qquad & \lambda_i^* = \beta \tfrac{\alpha^2 \sigma_e^2 + \sigma_\varepsilon^2}{2C\,(1-(1-\alpha)\overline{\lambda})^2} \end{aligned} \tag{2.14}$$

The latter equation is the best response function for player i, which gives the optimal choice for player i given the choice of the other players (the plot of this best response function for several different cost parameter values is given below in figure 2.1). It can clearly be seen that this optimal probability is increasing in the discount factor and the shock variances, and decreasing in the cost of information, which accords with what would intuitively be expected. The nature of the relationship between the individual optimum and the aggregate probability is given by:

$$\frac{\partial \lambda_i^*}{\partial \overline{\lambda}} = \frac{(1-\alpha)\beta\,(\alpha^2\sigma_e^2 + \sigma_\varepsilon^2)}{C\,(1-(1-\alpha)\overline{\lambda})^3} \tag{2.15}$$

Hence for $0 < \alpha < 1$ the optimal probability is increasing in the choice of the other agents, for $\alpha = 1$ the optimal individual probability does not depend on the strategy of other agents and for $\alpha > 1$ the optimal probability is decreasing in $\overline{\lambda}$. These three cases represent situations of strategic complementarity, strategic neutrality and strategic substitutability, respectively. The role of α in the original model can be seen from using equation 2.7 to replace output in equation 2.4. For $0 < \alpha < 1$ the optimal individual price depends positively on aggregate price whereas for $\alpha > 1$ the optimal individual price depends negatively on the aggregate price. This strategic relationship (which Mankiw and Reis (2002) note is referred to by Ball and Romer (1990) in relation to monetary non-neutrality) is inherited by the information updating 'game'. Hellwig and Veldkamp (2009) find that for a model in which the information choice is modelled as a signal extraction problem, rather than sticky information, the strategic relationships in the information choice are the same as those in the underlying economic problem.

Cooper (1999) shows that for payoff function maximization a positive cross-partial derivative with respect to strategies implies strategic complementarity; here the 'payoff function' is to be minimized - consequently a negative cross-partial derivative would indicate strategic complementarity:

$$\frac{\partial^2 Loss_i}{\partial \lambda_i \overline{\lambda}} = \frac{-2\,\beta\,(1-\alpha)\,(\alpha^2\sigma_e^2 + \sigma_\varepsilon^2)}{(1-(1-\alpha)\overline{\lambda})^3} \quad = \begin{cases} < 0 & if\; 0 < \alpha < 1 \\ = 0 & if\; \alpha = 1 \\ > 0 & if\; \alpha > 1 \end{cases} \tag{2.16}$$

It can be seen from the above that the different strategic relationships do indeed obtain for the different values of α as described in the discussion of the best response function. Therefore the actual value, or plausible range of values, of α will be crucial in the outcome of the game. Cooper and John (1988) show that under strategic substitutability there is a unique symmetric Nash equilibrium, whereas under strategic complementarity there may be multiple symmetric Nash equilibria. Graphically symmetric equilibria are given by the intersections of the best response function with the 45^o line, as these are the points at which $\lambda_i^* = \overline{\lambda}$[9]. For the best response function of equation 2.14 the second derivative with respect to $\overline{\lambda}$ is

$$\frac{\partial^2 \lambda_i^*}{\partial \overline{\lambda}^2} = \frac{3\,(-(1-\alpha))^2 \beta\,(\alpha^2\sigma_e^2 + \sigma_\varepsilon^2)}{(1-(1-\alpha)\overline{\lambda})^4} = \begin{cases} > 0 & if\ \alpha \neq 0 \\ = 0 & if\ \alpha = 0 \end{cases} \tag{2.17}$$

Therefore the best response function is convex in $\overline{\lambda}$ and consequently there can be a maximum of two symmetric Nash equilibria (i.e. two crossings of the 45^o line) although there may be fewer than this depending on the particular curvature of the best response function for given parameters. A caveat to this is that while the strategies of the players are restricted to $[0, 1]$ the best response function as given above does not necessarily map into this interval for all parameter values. This can give rise to an outcome of two 'interior' equilibria which satisfy the optimization and then a third equilibrium at $\lambda_i^* = \overline{\lambda} = 1$, which is an equilibrium in the sense that for all players it is the best possible strategy. There may be an outcome of this latter equilibrium alone if the 'true' best response function does not cross the 45^o line at all in the $[0, 1]$ interval. A possible solution to this is to restrict the parameter values such that equation 2.14 gives an optimal response of 1 for $\overline{\lambda} = 1$ in which case, provided that $\overline{\lambda} = 0$ does not give an optimal response of 0, by convexity of the best response function there will be no more than one symmetric equilibrium within the $(0, 1)$ interval. This restriction on the range of the best response function also has the advantage that it allows the standard game theory approach of using fixed point theorems to prove the existence of a symmetric equilibrium to be applied[10]. However, insofar as some parameter values were calibrated

[9] See for example figure 2.1

[10] Branch et al. (2009) define the best response function as the λ which minimizes the loss with the restriction $\lambda \in [0, 1]$, rather than explicitly deriving an equivalent of 2.14, and claim existence on the basis of Brouwer's fixed point theorem.

it might not be valid to restrict the non-calibrated parameters in order to achieve this. Furthermore for a multi-period problem it would typically not be possible to impose this restriction because it is not possible to derive an explicit algebraic representation of the best response function. Therefore the possibility of multiple interior equilibria will apply for the longer planning horizon.

This analysis of the one-period problem accords with the analysis by Branch et al. (2009), where it is shown that for a loss function based on the pricing forecast error for a general single period and some cost $C\lambda^2$, the loss function is monotonically decreasing in the update probability. It is therein claimed that when multiple equilibria occur they may be discriminated between according to stability, where stable is meant as the slope of the best response function being less than one, such that, for a small change in the aggregate strategy, the change in the optimal strategy is smaller and so the original symmetric strategy would be 'returned to' as a symmetric equilibrium.

Figure 2.1 shows several instances of the best response function given by equation 2.14 for values of C between 0 and 20, with the other relevant parameters set to $\beta = 0.99$, $\alpha = 0.1$, $\sigma_e^2 = 0.1$, $\sigma_\varepsilon^2 = 0.1$. The latter two values are chosen to match the shock variance used in Branch et al. (2009), the value of β is standard for a setting where one period is thought of as one quarter and the value of α is within the range considered reasonable given the definition of α (see appendix A.1.1). The horizontal black line in 2.1 represents the upper bound on λ, so the total best resposse function can be thought of as the red curve (showing the interior best responses) plus the black horizontal line for the remainder of the range of $\overline{\lambda}$, if the red curve reaches the upper bound prior to $\overline{\lambda} = 1$. It can be seen that for very low cost the only symmetric equilibrium is $\lambda = 1$; for intermediate cost there are two interior symmetric equilibria in addition to this and as cost increases these equilibria move towards the extremes, until for very high cost there is only a single symmetric equilibrium close to zero. Figure 2 of Branch et al. (2009) shows a similar relationship between information cost and the optimal updating probability (although the form of the best response functions there differs due to the additional policy parameters in that model). This plot is sufficient to represent the equilibria as the agents are identical, hence they all have the same best response function.

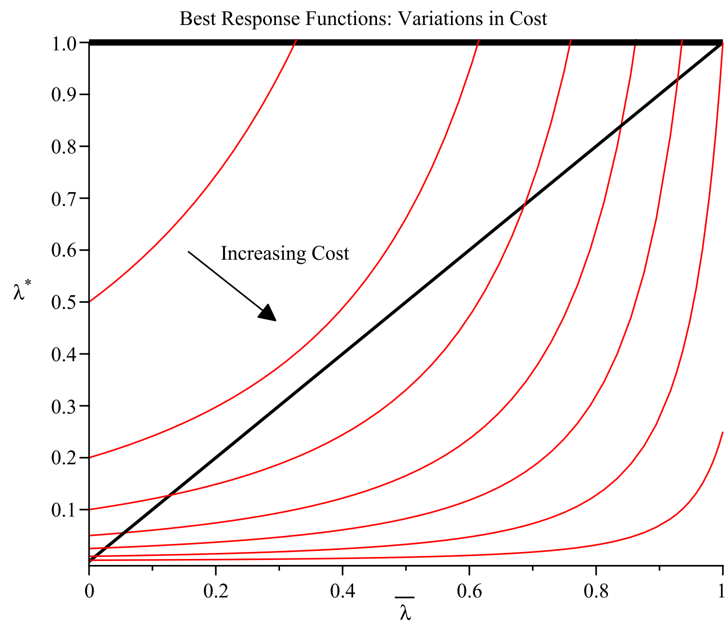

Figure 2.1: Best response functions for the one-period problem

The remainder of this chapter goes beyond Branch et al. (2009) in explicitly formulating a multi-period loss function with a definite starting date, rather than using the expected per-period loss under the asymptotic information distribution[11], and solving numerically for the optimal probability in order to see how the choice depends also on the length of the planning horizon, rather than determining the equilibrium probability which would arise in an arbitrary period in the distant future. The results of this are presented in the next section.

2.4 Multi-period Loss Function

For the multi-period loss function as shown in equation 2.11 and for the asymptotic loss function in equation 2.12 it is not possible to derive the exact best response functions, as discussed above, and consequently it is not possible to consider analytically the relationship between the aggregate updating probability and the individual updating

[11]The use of such a loss function in Branch et al. (2009) is a further reason why it may not be appropriate to use the results therein to explain empirical dynamics.

probability. However, using again the condition from Cooper (1999) concerning the sign of the cross-partial derivative of the loss function with respect to strategies it is possible to conclude that the value of the parameter α will continue to determine whether the game features strategic complementarity or substitutability (see appendix A.1.3 for details). Specifically, it will continue to be the case that strategic complementarity in pricing ($\alpha < 1$) results in strategic complementarity in choice of updating probability and strategic substitutability in pricing ($\alpha > 1$) results in strategic substitutability in choice of updating probability. Therefore for the multi-period loss function the value of α will continue to determine whether or not multiple symmetric equilibria are possible.

The remainder of this section presents the results of numerically searching (with $\overline{\lambda}$ grid increments of 0.005[12]) for the symmetric Nash equilibrium probabilities with varying parameter values[13]. Typically, for the horizon of the loss function in equation 2.10 set to 15 periods the minimizations in relation to variations in a single parameter take approximately one hour. In the case of the varying parameter being N the largest value used was 50, in which case the computation took approximately eight hours.

Individual parameter variation will indicate the effect of each parameter on the equilibria, such as whether the number of equilibria changes or the equilibria respond monotonically. As the purpose of this is simply to see the effects that parameter changes have, rather than being an attempt to produce results which could necessarily be claimed to be 'realistic', some parameter values are used which would be unlikely to feature if the model were calibrated[14]. The computations are undertaken for both of the loss functions detailed above, in order to distinguish whether any differences in the results in comparison to the existing literature are due to the different specification of the loss function or to the exclusion of an optimizing policy-maker. Generally it is found that the best response functions do not vary qualitatively between the two loss functions and that the

[12] As will be seen in the plots below, the resulting best response functions are sufficiently regular that the increase in precision obtained by using a finer grid would be unlikely to have much economic significance.

[13] The code for this is written in Python and uses the fminbound minimization routine from the the scipy.optimize package

[14] For example, very large shock variances will be investigated. Given that the cost parameter as specified here is a free parameter, a value could always be chosen such that the resulting λ^* would match equivalent empirical estimates, such as Mankiw and Reis (2007).

quantitative differences are not large. Only the results for the loss function as in equation 2.10 are presented below, therefore. While a direct comparison with Branch et al. (2009) is not possible for these best response functions, due to its limited presentation of such results, the results here do suggest that any difference between the two models in terms of macroeconomic outcomes may be the result of the difference in policy rather than the specification of the loss. Given the non-linearity of the loss function with respect to most of the parameters it may be worthwhile to consider simultaneous variation of several parameters. However due to the number of possible variations this would imply it might be impractical to systematically investigate this.

2.4.1 α

The best response functions for variations in α are shown in figure 2.2. The other parameters were set to $\beta = 0.99$, $\sigma_e^2 = 0.2$, $\sigma_\varepsilon^2 = 0.2$, $\delta = 0.6$, $\rho = 0.6$, $C = 5$ *and* $N = 15$. The range of values of α used is 0.01 - 2.

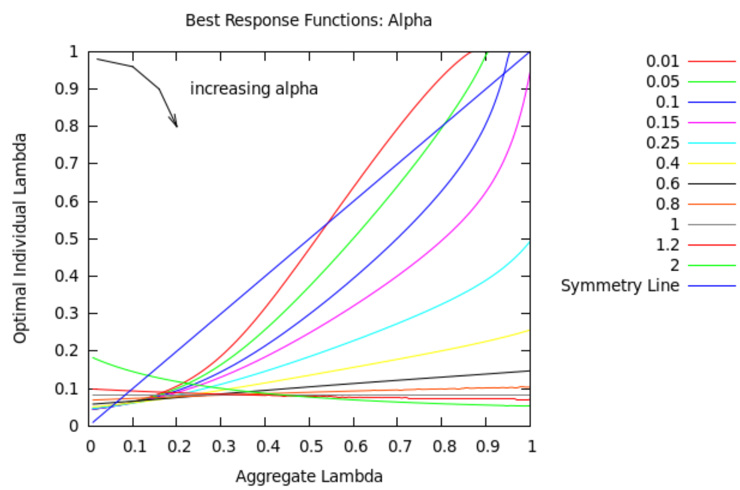

Figure 2.2: Best response functions for $0.01 \leq \alpha \leq 2$

The effect of parameter variation here appears to be a rotation of the best response function (with perhaps some change in curvature also), rather than approximately a translation as in the previous case, with an increase in α reducing the slope of the best response function. In particular, the final two indicated best response functions, which correspond to values of α equal to 1.2 and 2, are negatively sloped. The indicated best response function for $\alpha = 1$ has zero slope and the remaining curves are all for values of $0 < \alpha < 1$. This confirms the earlier analysis[15] that the value of α determines the nature of the information 'game' with respect to strategic substitutability/complementarity.

A further implication of this effect of α is that the lower is the degree of strategic complementarity (i.e. the greater is α) the less likely is it that there will be multiple interior equilibria. A possible explanation of this is that for $0 < \alpha < 1$ the best response function is relatively shallow-sloped for low values of $\overline{\lambda}$ either because rivals who update more frequently (on average) attach a high weight to the prices set in previous periods[16] (for values of α closer to 0 - 'strong complementarity'), such that those who do not receive new information will not experience large relative price differentials, or because having rivals update more frequently on average does not imply a large loss to the agent relative to the cost of increasing λ_i (for values of α closer to 1 - 'weak complementarity'), so the incentive to have an updating probability close to that of other agents is small. For higher values of $\overline{\lambda}$ and lower values of α the best response function becomes steeply-sloped because now those agents who update on average more often will not attach a large weight to prices set in previous periods, so those agents who do not receive new information will be 'left behind' and experience relatively large deviations from the optimal price. Therefore over this range agents have more incentive to choose an updating probability close to that of other agents, and in some cases higher than that of other agents.

2.4.2 Cost

Figure 2.3 shows the calculated best response functions for varying values of C, with other parameters constant at $\beta = 0.99$, $\alpha = 0.1$, $\sigma_e^2 = 0.2$, $\sigma_\varepsilon^2 = 0.2$, $\delta = 0.6$, $\rho = 0.6$ *and* $N = 15$. These values are again not calibrated, but chosen to give intermediate results in the sense of a varying number of equilibria of different values. The 45 degree line is also shown, the intersection of which by the best response functions indicates the symmetric equilibria, and the upper limit of 1 is also marked. The range of values of C

[15] Based on the condition from Cooper (1999) concerning the cross-partial derivative of the loss function with respect to strategies.

[16] Equivalently, the weight on new information in the optimal price is small.

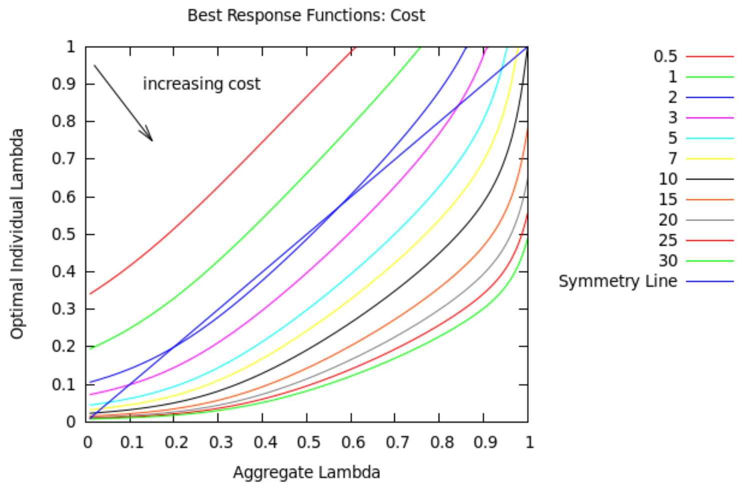

Figure 2.3: Best response functions for $0.5 \leq \mathrm{C} \leq 30$

is 0.5 - 30 (the legend shows the exact value for each function).

It can be seen that for very low cost there are no symmetric equilibria below $\lambda = 1$. As cost increases two interior equilibria occur, with further increases in cost reducing the lower of these and increasing the higher of these. While it may seem contrary to the intuition of the model that higher cost could lead to a higher equilibrium rate of information acquisition, this is a consequence of the curvature of the best response functions (see discussion above concerning the slope of the best response function), and it remains the case that for any given aggregate probability the optimal individual probability is lower as cost increases, so at higher cost it is not optimal to match the choice of other agents until the weight on new information in the optimal price (i.e. $\overline{\lambda}$) is higher. It can also be seen that of the two interior symmetric equilibria it is the lower equilibrium which is stable in the sense discussed above in relation to Branch et al. (2009), so it could be argued that this suggests that, in a properly dynamic model, from either equilibrium for a given level of cost, an increase in cost would actually lead to a lower equilibrium updating probability.

2.4.3 N

Figure 2.4 shows the best response functions for variations in N with other parameters set to $\beta = 0.99,\ \alpha = 0.1,\ \sigma_e^2 = 0.2,\ \sigma_\varepsilon^2 = 0.2,\ \delta = 0.9,\ C = 2\ and\ \rho = 0.8$. The range of values of N used is 2 - 50. The second and third plots in figure 2.4 show the lower and upper equilibria with the scale adjusted such the differences are discernible. It can be seen that the lower equilibrium is increasing in N and the upper equilibrium is decreasing in N. The explanation for this is that because the total expectational loss is increasing faster than the total acquisition cost[17] as N increases, the range of λ values for which it is optimal to choose $\lambda_i < \overline{\lambda}$ is smaller. It may be expected that the longer is the planning horizon the greater are the potential losses resulting from not receiving new information, so as N increases eventually the only equilibrium will be $\lambda^* = 1$. However, it can be shown that the loss functions are convergent as $N \to \infty$ (see appendix A.1.4), so there will always be some possible parameterization such that at least one interior equilibrium exists[18] (although whether this parameterization was appropriate would be an empirical issue).

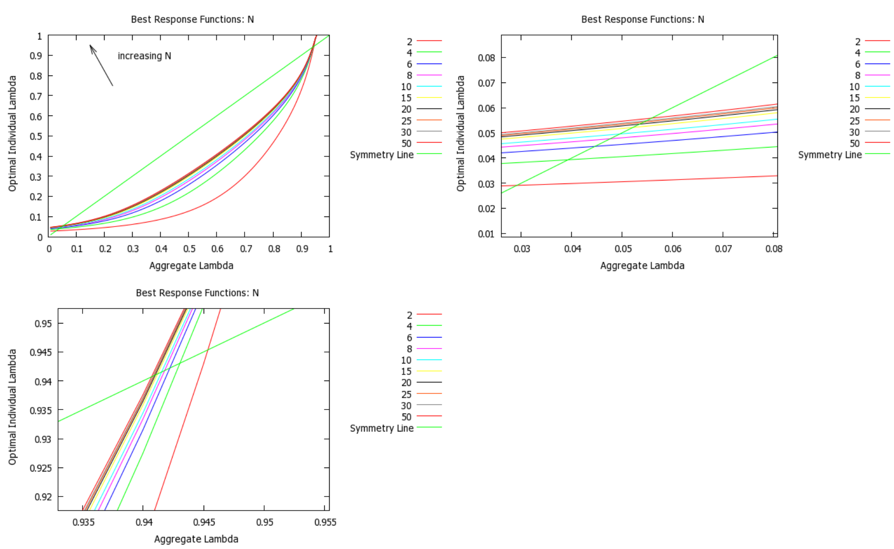

Figure 2.4: Best response functions for $2 \leq N \leq 45$

[17]The expected price deviation is increasing with N, while C is constant.

[18]The parameterization used here appears to be close to giving invariant equilibria as N varies.

2.4.4 ρ

Figure 2.5 shows the best response functions for variations in ρ, with other parameters set to $\beta = 0.99,\ \alpha = 0.1,\ \sigma_e^2 = 0.2,\ \sigma_\varepsilon^2 = 0.2,\ \delta = 0.6,\ C = 5\ and\ N = 15$. The range of values of ρ used is 0 - 0.999.

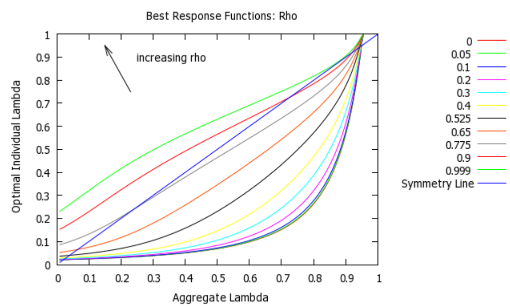

Figure 2.5: Best response functions for $0 \leq \rho \leq 0.99$

It can be seen that as ρ (the persistence of disturbances to the optimal price) increases there are at first only small changes in the equilibria, with the lower (stable) equilibrium increasing slightly and the upper (unstable) equilibrium decreasing slightly. Again the behaviour of the lower equilibrium is what might be expected, in that the more persistent are shocks the greater may be the cost of not knowing their realizations. As ρ increases the interior equilibria are eliminated and $\lambda = 1$ is the only symmetric equilibrium.

2.4.5 δ

Figure 2.6 shows the best response functions for variations in δ, with other parameters set to $\beta = 0.99,\ \alpha = 0.8,\ \sigma_e^2 = 0.2,\ \sigma_\varepsilon^2 = 0.2,\ \rho = 0.8,\ C = 5\ and\ N = 15$. The range of values of δ used is 0 - 0.99.

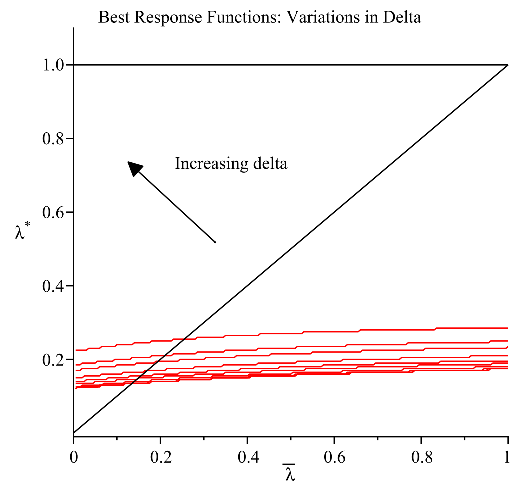

Figure 2.6: Best response functions for $0 \leq \delta \leq 0.99$

Here a much larger value for α has been used than in the previous cases in this section. This is because, as can be seen from equation 2.11, the δ terms are multiplied by α^2, so in order for the changes in δ to have a discernible effect on the best response functions it is necessary to use a relatively high value of α. As a consequence of this the best response functions are near-linear, with only a single symmetric equilibirum in each case. From figure 2.6 it can be seen that increasing δ increases the equilibrium λ, the implication of which is that the greater is the persistence of disturbances to nominal demand the greater is the equilibrium updating probability.

2.4.6 σ_e^2

The best response functions for changes in σ_e^2 are shown in figure 2.7, for other parameters set to $\beta = 0.99,\ \alpha = 0.1,\ \sigma_\varepsilon^2 = 0.2,\ \delta = 0.6,\ \rho = 0.6,\ C = 5\ and\ N = 15$.

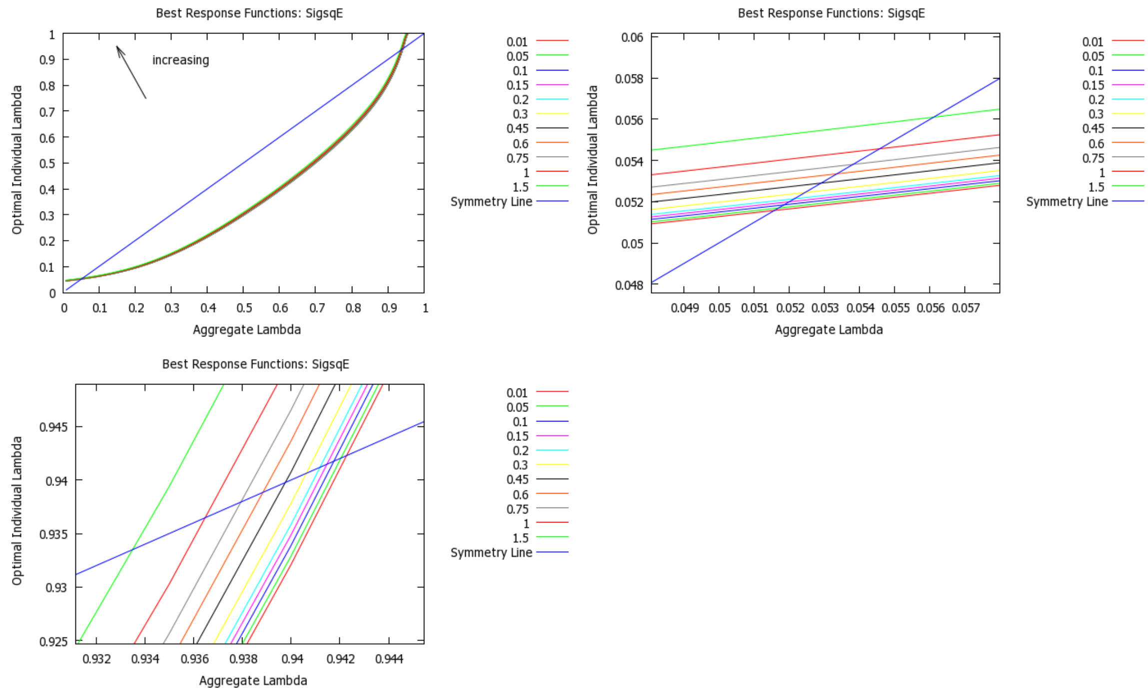

Figure 2.7: Best response functions for $0.01 \leq \sigma_e^2 \leq 1.5$

It can be seen from figure 2.7, with the second and third plots showing the intersection points with rescaled axes, that the optimal response is increasing in the variance of demand shocks, which accords with the forecast error of non-updaters being increasing in the shock variance. Again, the symmetric equilibria are increasing at low values of λ and decreasing at high values. However the extent to which λ might change in the event of a realistic change in σ_e^2 may not be very great[19], given that empirical estimates of the variance of shocks to nominal demand are at the lower end (or lower) of the range used here[20], at least in cases of strategic complementarity.

2.4.7 σ_ε^2

Figure 2.8 shows the best response functions for varying σ_ε^2 with other parameters set to $\beta = 0.99$, $\alpha = 0.1$, $\sigma_e^2 = 0.2$, $\delta = 0.6$, $\rho = 0.6$, $C = 5$ *and* $N = 15$. It can be

[19] As this term is weighted in the loss function by α^2.

[20] E.g. Mankiw and Reis (2002) estimate a similar process using USA data and find a value of 0.007 for the standard deviation

seen that as σ_ε^2 increases the stable equilibrium probability is increasing, which again is in accordance with the intuition of the agents' incentives to update. The changes in the equilibrium probability are relatively large for variations in this parameter, although again it may be that the actual variance of such a disturbance would be at the lower extreme of the range of values considered.

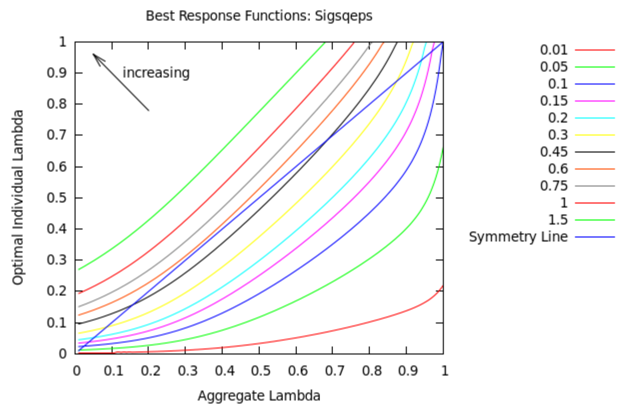

Figure 2.8: Best response functions for $0.01 \leq \sigma_\varepsilon^2 \leq 1.5$

Taken together, the results of sections 2.4.4 - 2.4.7 suggest that the equilibrium information updating probability is more sensitive to changes in the process of the (aggregate) mark-up shock to the optimal price than to changes in the process of the shock to nominal aggregate demand. However, as the discussion above indicates, this depends crucially on the value of α; if $\alpha > 1$ (strategic substitutability) then equation 2.11 suggests that this would be reversed, with the equilibrium updating probability being more responsive to changes in the process of the demand disturbance. Given that in this model α depends on the underlying preference parameters and consumption substitution elasticity it is known that the strategic relationship between price-setters is one of complementarity, therefore of the two shocks it is the nature of the price shock to which the choice of updating probability is most responsive. Mackowiak and Wiederholt (2009) shows that for a rational inattention model in which price-setters face aggregate demand shocks and idiosyncratic price shocks, attention will be predominantly allocated to whichever shock is most volatile/has greatest effect on profits. A comparable result here is that, for $\alpha = 1$

and $\rho = \delta$, the equilibrium λ will be most sensitive to whichever shock has the greatest volatility.

2.4.8 β

Figure 2.9 shows the best response functions for variations in β with other parameters set to $\alpha = 0.1$, $\sigma_e^2 = 0.2$, $\sigma_\varepsilon^2 = 0.2$, $\delta = 0.9$, $\rho = 0.8$, $C = 2$ *and* $N = 15$. The range of β values used is 0.1 - 0.999. It can be seen that as β increases the 'stable' equilibrium λ increases. As the discount factor increases the consequence of future losses being less discounted is that the value of information in future is greater, so the optimal updating probability increases.

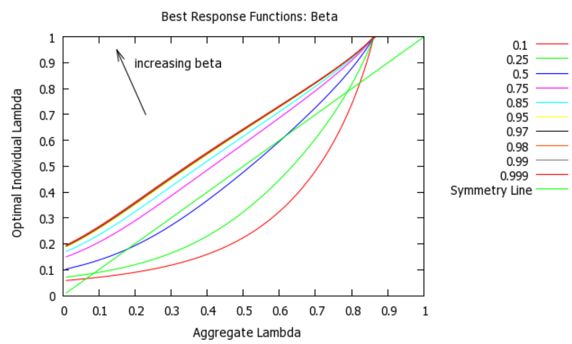

Figure 2.9: Best response functions for $0.5 \leq \beta \leq 0.99$

2.5 Macroeconomic Implications of Endogenous Updating

Having analyzed in the above section the endogenous updating probability under various parameterizations, the comparative static implications of this for the variances of output and inflation are now considered. In particular, the changes due to varying values of σ_e^2 are investigated. The variance of the money/nominal demand shock is the focus because this parameter is perhaps the closest to a policy variable in this model. For instance, if equation 2.8 was a policy rule then increased competence or determination to achieve the target could be expressed as a reduction in σ_e^2. The following compares the outcomes of such a change under exogenous and endogenous sticky information.

Firstly, figure 2.10 shows the variance of output as σ_e^2 varies from 0 to 1.6. The figure shows two plots because for the parameterization used (which is the same as subsection 2.4.6) there are two interior symmetric equilibrium λ values. In both cases the value of λ in the exogenous case is set equal to the endogenous λ when $\sigma_e^2 = 0.2$. The left-hand plot shows the variance for the lower equilibrium value of λ and the right-hand plot the variance for the higher equilibrium value of λ[21]. The points in the plot are not connected because in the absence of an equilibrium selection mechanism (such as, in a dynamic determination, stability) it is not possible to rule out a switch between equilibria as the shock variance varies. Figure 2.11 shows the same plots for the variance of inflation. In contrast to Branch et al. (2009), the finding from these two figures is that there is no trade-off between output and inflation variance in the exogenous λ case as σ_e^2 varies, because both variables are increasing in σ_e^2 (this is the direct effect of σ_e^2) , whereas with endogenous λ there is a trade-off due to the indirect effect of σ_e^2, which is the effect of σ_e^2 on λ. This is because the output variance is decreasing in λ while the inflation variance is increasing in λ, therefore a change in σ_e^2 which increases (decreases) the equilibrium λ will decrease (increase) the output variance and increase (decrease) the inflation variance, as shown in figures 2.10 and 2.11. The explanation for this contradiction is perhaps most likely to be that these results consider a change in the variance of a fundamental shock, which has at best a pseudo-policy interpretation, whereas Branch et al. (2009) use an explicit, restrictive representation of policy which has further effects on the choice of λ.

[21]It appears that the variance in both cases coincides in the right-hand plot. It does not, but the difference between the two is too small to plot collectively.

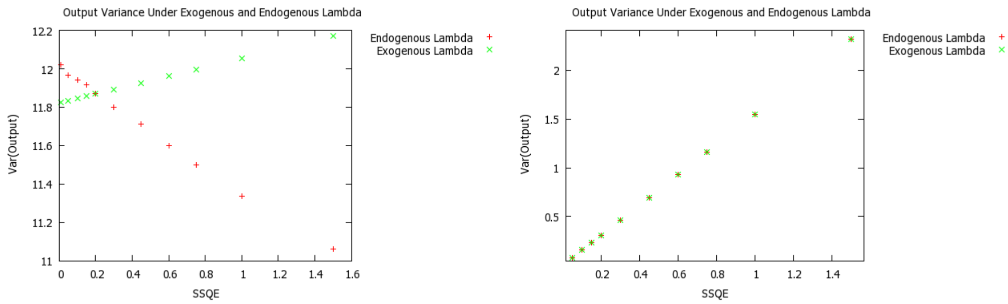

Figure 2.10: Output Variances Given Multiple Symmetric Equilibria

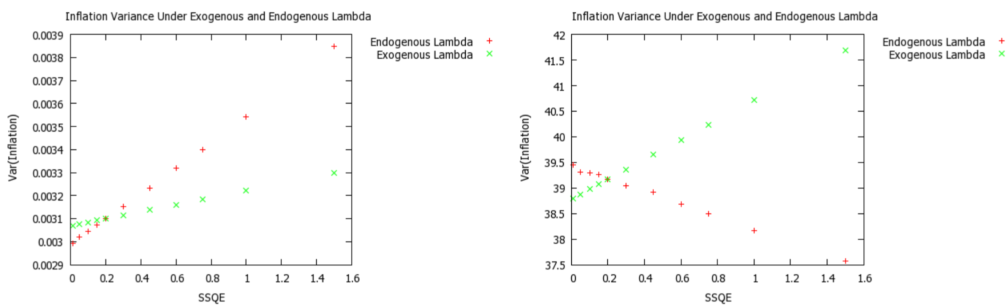

Figure 2.11: Inflation Variances Given Multiple Symmetric Equilibria

2.6 Conclusion

One potentially interesting extension of this model would be to allow the shocks to be idiosyncratic - above both ε_t and e_t are aggregate shocks, i.e. common to all agents. If the model featured idiosyncratic pricing shocks, say, then it might be expected that agents would choose different updating probabilities, although this effect may be limited by the strategic complementarity in price-setting, as discussed above. Similarly, if agents had differential updating probabilities with respect to the different shocks (as in Mackowiak and Wiederholt (2009) then this might imply that the responses to the different shocks, and to changes in the nature of the shocks, would be fundamentally different, apart from any differences arising from the direct effect of the shock. As yet this type of model has used the rational inattention framework rather than the sticky information model, but this could be attempted and it might be expected to give particularly interesting results in the model of chapter 3, in which there are different types of agents also, so it could be the case that agents would 'focus' on the shock most directly applicable to them rather than all shocks equally. Such a model could be one course of further development of this topic.

The results above demonstrate that it is possible to endogenize the determination of the Mankiw-Reis information friction in a simple, general model and that the relationship between this determination and the parameters of the model is as would be predicted in accordance with economic logic. Attempts have been made within the existing literature to apply such a framework empirically, both with exogenous updating (e.g. Coibion (2010) and Reis (2009)) and, to a limited extent, with endogenous updating (Branch et al. (2009)). However, the difficulty with such applications is that either the cost is a free parameter, and so can be adjusted to achieve a desired updating probability, or the estimated probability is difficult to reject, in the sense that a wide range of values is inherently plausible and available microeconomic evidence concerns the frequency of actual price changes rather than that of price plan changes. Consequently, while the sticky information model has theoretical appeal it remains unclear whether its empirical performance is superior to alternatives such as the standard New Keynesian sticky price model. Those papers which do attempt empirical comparison, such as Coibion and Gorodnichenko (2011) or Paustian and Pytlarczyk (2006), assume exogenous determination of λ and so could be argued to provide inadequate grounds for making definitive conclusions concerning this[22]. The incorporation of endogenous determination of the updating probability may be one of the improvements necessary to enable the resolution of this issue. Furthermore, the results suggest that in order to correctly assess the

[22] Chapter 4 will discuss in more detail the literature concerning empirical assessment of sticky information models. This literature is now quite broad in terms of both models and estimation methods.

consequences of endogeneity for macroeconomic modelling and policy it is necessary to have as realistic a representation of the determination of information frictions (and the effect on this of the implementation of policy) as possible.

Chapter 3

Endogenous Information Updating in a DSGE Model

3.1 Introduction

As seen in chapter 2, the early implementations of the 'sticky information' concept used relatively simple macroeconomic models to illustrate, unobscured by other factors, the consequences of this type of imperfect information. However, a possible drawback of this is that such models do not include all economically important components. Therefore, later work has extended this by incorporating the informational friction into more detailed models featuring a variety of agents. In particular, this chapter will focus on Mankiw and Reis (2006b), who construct a DSGE model in which distinct firms, consumers and workers are potentially inattentive[1]. This is modelled as all three agent types being subject to 'sticky information', with the result that prices, nominal wages and consumption plans are all based on out-of-date information, to some extent. While these are not rigidities in the sense of variables being fixed at constant for several periods (because prices, wages and consumption can all be costlessly varied in each period), the effect is analogous in that the economy may not immediately respond fully to shocks,

[1] The material in the article was later published as Mankiw and Reis (2007), however this chapter refers to the working paper because it contains a more detailed description of the model and the same results as in the published piece.

so the information restrictions can still be thought of as a form of nominal (and in the case of consumption, real) rigidity. This also raises the possibility that when updating probabilities are endogenized there will not only be strategic interaction between agents of the same type but also between agents of different types, as they will each be affected by the others' choices. That such a model is more appropriate than the earlier model with a single type of agent is supported by the results of Mankiw and Reis (2006a), who find that the ability of an estimated model to fit the data is improved considerably when all three agent types, rather than only a proper subset thereof, are inattentive. In particular, they find that such a model is able to replicate the positive correlation between changes in inflation and the output gap, the change in real wages being proportionally smaller than the change in productivity and delayed maximal responses to exogenous shocks. To some extent the model with only a single type of inattentive agent is able to fit some of these regularities, especially the latter, but to fit all three quite well the more comprehensive model is necessary.

An attempt to incorporate endogenous updating probabilities in an essentially similar model is made by Drager (2011). The specification therein features several differences from the model here and these differences will be discussed below once the model has been presented. The general conclusion of Drager (2011) is that such a model is capable of producing persistent output and inflation and 'hump-shaped' responses of output and inflation to monetary policy shocks, purely on the basis of the information frictions.

It is found here that the optimal updating probabilities depend upon several factors. As may have been expected, the optimal probabilities decrease as the cost associated with receiving new information and updating plans increases. The model exhibits strategic inter-relationships between different agents. As in chapter 2, agents of the same type generally have a positive effect on each other's choice of updating probability - this effect is especially strong in the case of price-setting firms. However, it is found that between agents of different types there is an asymmetric effect, with firms reacting quite strongly to the choice of households but households being relatively indifferent to the updating choice of firms. The responsiveness of monetary policy is also found to significantly affect the degree of inattentiveness. If monetary policy changes to react more strongly to inflation then this not only reduces inflation variability directly; there is also an indirect effect from the induced reduction in updating probabilities (as private agents needn't pay as much attention if policy causes greater stability), such that this policy is more effective under endogenous sticky information than under exogenous sticky information, although the 'trade-off' for this is that the reduction in output volatility is slightly less than would occur under exogenous sticky information. A similar indirect amplification is found to occur when the variability of the policy disturbance changes. A further finding in this chapter is that policy changes and the resulting changes in updating probabilities alter the dynamic responses of the economy, in terms of both the size of the responses and, to a lesser extent, the time profile, i.e. the length of time taken for the maximal response to a shock to occur. One of the main findings of chapter 2 was that multiple updating equilibria were possible. Here the potential for multiple equilibria appears to be much less than was the case for the relatively simple model studied in chapter 2, although it is not possible to investigate this as extensively numerically due to the presence of agents of different types.

In order to consider the model in detail section 3.2 reproduces the relevant equations from Mankiw and Reis (2006b)[2], with the solution to their model then being used to calculate the loss due to being imperfectly informed, in section 3.3. The computation of the model solution and simulations is discussed in section 3.4, then the numerical results are presented in section 3.5. Section 3.6 concludes this chapter with a summary of the results.

3.2 Model

The full statement of the model, including all derivations and an explanation of the solution method, is given in the appendix. The following outlines the initial specification and highlights those parts which are significantly affected by the information restrictions.

The agents featured are households (sub-divided into consumers, shoppers and workers), firms (sub-divided into a pricing department and hiring department) and a policy-maker. The model is generally of the New Keynesian, monopolistic competition type, with the substitution of the information rigidity in place of other nominal rigidities. For

[2] The full derivation of the model is given in appendix A.2.2.

instance, firms can set new prices in each period. Sub-division of households and firms is necessitated by the desire to have only one type of inattentive agent in any one particular market, in order to permit the exclusion of possible strategic decisions relating to other agents' inattentiveness and in order to ensure that there are no incompatibilities between the choices of 'sub-agents' within an entity. An additional motivation for this sub-division is that it ensures that the prices of individual varieties of goods and labour are correctly and uniformly perceived by their purchasers, which will simplify aggregation across agents.

The actions of the government/policy-maker are to set monetary policy, described by a Taylor rule:

$$I_t = \bar{I}\left(\frac{Y_t}{Y_t^n}\right)^{\phi_y}\left(\frac{P_t}{P_{t-1}}\right)^{\phi_p} e^{-\varepsilon_t} \tag{3.1}$$

and to demand some proportion of aggregate output, given by $1 - 1/G_t$, where G_t is a stochastic process, defined below. This demand can be thought of as the realization of fiscal policy. The rule in equation 3.1 does not feature the nominal interest rate in the previous period, i.e. there is no interest rate smoothing; instead the shocks to monetary policy are assumed to be AR(1) (in log form), with the process implicitly described in equation 3.20.

The disaggregated goods market features shoppers with full price information and pricing departments with partial information about aggregate variables, as a consequence of receiving new information in each period with probability λ. In period t the shopper of household h solves[3]:

$$min_{\{C_{t,h}(f)\}_{f\in[0,1]}} \int_0^1 P_{t,f}C_{t,h}(f)\, df \; s.t. \; C_{t,h} = \left(\int_0^1 C_{t,h}(f)^{\frac{\nu_t-1}{\nu_t}} df\right)^{\frac{\nu_t}{\nu_t-1}} \tag{3.2}$$

where the constraint is the combination of goods, indexed by f, to give composite consumption by household h. The solution gives demand for good f by household h as:

[3]Throughout households are indexed by h and firms by f.

$$C_{t,h}(f) = C_{t,h} \left(\frac{P_t(f)}{P_t} \right)^{-\nu_t} \tag{3.3}$$

As the shopper is assumed to be fully informed about prices the composition of the consumption bundle, in the sense of the share of each good in total consumption, will be identical across households. As a result of this, the quantities of this bundle chosen by imperfectly informed consumers are easily aggregated.

The labour market features hiring departments of firms, which are fully-informed about nominal wages, and households (consumers/workers), who are partially-informed about aggregate variables, as a consequence of their information set being updated with probability ω in each period. Nominal wages are set by households, which can be thought of as having this wage-setting power due to the presence of (otherwise non-modelled) trade unions and due to labour type being varied across households in the same way that goods produced by firms are considered to be continually various. The hiring department seeks to minimize the cost of hiring a unit of composite labour, given the labour composition constraint:

$$min_{\{L_{t,f}(h)\}_{h\in[0,1]}} \int_0^1 W_{t,h} L_{t,f}(h)\, dh \; s.t. \; L_{t,f} = \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} \tag{3.4}$$

The resulting labour demand function is:

$$L_{t,f}(h) = L_{t,f} \left(\frac{W_t(h)}{W_t} \right)^{-\gamma_t} \tag{3.5}$$

It can be seen from equations 3.3 and 3.5 that the agents who are not subject to information restrictions behave as in typical New Keynesian models with two-stage budgeting.

Households are assumed to all have the same preferences, given by:

$$U(C_{t,h},\ L_{t,h}) = ln\Big(C_{t,h}\Big) - \frac{L_{t,h}^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \tag{3.6}$$

The lifetime utility maximization objective is therefore:

$$\sum_{s=0}^{\infty} \xi^s \left(ln\Big(C_{t+s,h}\Big) - \frac{L_{t+s,h}^{1+\frac{1}{\psi}} + 1}{1+\frac{1}{\psi}} \right) \tag{3.7}$$

In models with fully informed consumers consumption in each period is set based on the forecast of a function of future consumption, inflation and the nominal interest rate (equivalently, future consumption and future real interest rate). Here consumer behaviour is the same, except that the forecast may now be based on information which was last updated several periods ago. Over all agents there will be a distribution of information of different ages, so different households will choose different paths for the quantity of consumption, even though none of the shocks are idiosyncratic. Therefore aggregate consumption is an index over all households, as in equation 5.85 in appendix A.2.2 (in the log-linearized case). The model of Mankiw and Reis (2006b) allows a different updating probability between consumers and workers, however this is replaced here by a single updating probability which is applicable to households with respect to wage and consumption decisions.

Maximization of lifetime utility subject to budget and labour demand constraints gives (after re-arrangement of the relevant first-order conditions):

$$E_{t-k}\Big[C_{t,h}\Big] \equiv C_{t,h}^{k} = \frac{1}{\xi} E_{t-k}\left[\frac{C_{t+1,h}P_{t+1}}{I_t P_t}\right] \tag{3.8}$$

and

$$E_{t-k}\Big[W_{t,h}\Big] \equiv W_{t,h}^{k} = E_{t-k}\left[\frac{\gamma_t}{\gamma_t - 1} C_{t,h} P_t L_{t,h}^{\frac{1}{\psi}}\right] \tag{3.9}$$

where equation 3.8 gives the choice for consumption and equation 3.9 gives the choice for the nominal wage. ξ is the discount factor and I_t is the gross nominal one-period interest rate. These equations show the effect of the information rigidity, as both of the choice variables are now set based on expectations conditional on an information set of age k, rather than necessarily all information up to and including period t.

It can be seen from equation 3.9 that if households had full information then the nominal wage would not depend upon any expectation of future variables, only the contemporary values of the mark-up, consumption expenditure and the marginal disutility of work. However, now households with out-of-date information must set the nominal wage according to their conditional expectation of this. As a result nominal wages will vary across households, despite mark-up shocks being common to all.

The pricing departments of firms seek to maximize profits subject to the constraints of the production function and demand for their products. The profit function is:

$$\pi_{t,f} = \frac{P_{t,f}Y_{t,f}}{P_t} - \frac{W_t L_{t,f}}{P_t} \tag{3.10}$$

and the production function is:

$$Y_{t,f} = A_t L_{t,f}^{\beta} \tag{3.11}$$

where A_t is a stochastic technology variable, the process for which is implicitly specified in equation 3.21. Maximization of profits subject to the constraints leads to the pricing rule:

$$E_{t-j}\left[P_{t,f}\right] \equiv P_{t,f}^{j} = E_{t-j}\left[\frac{\nu_t W_t L_{t,f}}{(\nu_t - 1)\beta Y_{t,f}}\right] \tag{3.12}$$

Here the age of the information held by the pricing department is denoted by j. If j were zero then the optimal price would be a function of the contemporary mark-up and real marginal cost. However, for non-zero j the firm has to set price according to the conditional expectation of these variables, i.e. an expected mark-up over expected real marginal cost. Thus there will be a distribution of prices over firms which will depend upon the information updating probability λ.

Given this model, log-linearization around the non-stochastic, zero-inflation steady state gives the following counterparts to the key equations above, where a $\sim$ denotes the deviation of the log of a variable from its steady state value. [4]:

$$\tilde{i}_t = \phi_y(\tilde{y}_t - \tilde{y}^n_t) + \phi_p(\tilde{p}_t - \tilde{p}_{t-1}) - \varepsilon_t \tag{3.13}$$

$$\tilde{c}^k_{t,h} \approx E_{t-k}\left[\tilde{c}_{t+1,h} + \tilde{p}_{t+1} - \tilde{p}_t - \tilde{i}_t\right] \tag{3.14}$$

$$\tilde{w}^k_{t,h} \approx E_{t-k}\left[\frac{-\tilde{\gamma}_t}{\bar{\gamma}-1} + \tilde{c}_{t,h} + \tilde{p}_t + \frac{\tilde{l}_{t,h}}{\psi}\right] \tag{3.15}$$

$$\tilde{p}^j_t \approx E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t + \tilde{l}_{t,f} - \tilde{y}_{t,f}\right] \tag{3.16}$$

In addition, the aggregates over agents are:

$$\tilde{c}_t = \omega\sum_{k=0}^{\infty}(1-\omega)^k \tilde{c}_t{}^k \tag{3.17}$$

$$\tilde{w}_t = \omega\sum_{k=0}^{\infty}(1-\omega)^k \tilde{w}^k_t \tag{3.18}$$

$$\tilde{p}_t = \lambda\sum_{k=0}^{\infty}(1-\lambda)^k \tilde{p}^k_t \tag{3.19}$$

The exogenous shock processes are specified such that[5]:

[4]For the derivation of the full set of log-linear equations see the appendix, section A.2.2.

[5]The equations up to and including 3.12 specify ν_t and γ_t as levels, but in the following $\tilde{\nu_t}$ and $\tilde{\gamma_t}$ refer to the log deviations of these variables.

$$\tilde{\varepsilon}_t = \rho_\varepsilon \tilde{\varepsilon}_{t-1} + e_t^\varepsilon \tag{3.20}$$

$$\Delta\tilde{a}_t = \rho_{\Delta a} \Delta\tilde{a}_{t-1} + e_t^{\Delta a} \tag{3.21}$$

$$\tilde{g}_t = \rho_g \tilde{g}_{t-1} + e_t^g \tag{3.22}$$

$$\tilde{\nu}_t = \rho_\nu \tilde{\nu}_{t-1} + e_t^\nu \tag{3.23}$$

$$\tilde{\gamma}_t = \rho_\gamma \tilde{\gamma}_{t-1} + e_t^\gamma \tag{3.24}$$

The error terms e_t are zero mean with constant variance and are mutually uncorrelated.

The model can be solved using the method of undetermined coefficients[6], which gives as the solution to the reduced form of the model:

$$z_t = \sum_s \sum_{n=0}^{\infty} \hat{z}_n(s) e_{t-n}^s \tag{3.25}$$

for $z \in \{\tilde{y},\, \tilde{p},\, \tilde{w},\, \tilde{l},\, \tilde{i}\}$ and $s \in \{\Delta a,\, g,\, \nu,\, \gamma,\, \varepsilon\}$, such that $\hat{z}_n(s)$ is the effect of disturbance s (n periods ago) on variable z at t. In practice there is a limit to the number of coefficients for which a solution can be found. However, this limit may not significantly affect the results provided that it is not too small; for example, Mankiw and Reis (2006b) solve for $\infty \approx N = 1000$ and find that the terms omitted relative to a longer solution (e.g. the responses to shocks 1001 periods ago, 1002 periods ago, etc.) are extremely small[7]. Alternatively, an extended solution method which combines the undetermined coefficients approach with the generalized Schur decomposition method (which is often used in the solution of linear systems without any lagged expectations) is proposed by Meyer-Gohde (2010)[8]. This has the advantage of being able to explicitly determine the number of lagged expectations which are significant, rather than arbitrarily imposing

[6] p18 Mankiw and Reis (2006b).

[7] ibid. p21.

[8] The MATLAB code to implement the Meyer-Gohde (2010) method is provided by the author at http://ideas.repec.org/c/dge/qmrbcd/171.html

a truncation. It is shown[9] that such a truncation may considerably alter the results, comparing solutions for N = 10 and N = 100. However, it is not clear that the proposed method would be significantly superior to a truncation with larger N. The method of Meyer-Gohde (2010) produces a solution of the same form as equation 3.25, but is applied to the full set of endogenous variables, rather than those of the reduced form of the model. Hence this method may be superior to that of Mankiw and Reis (2006b) when applied generally; however it may not offer an improvement on the particular solution to the Mankiw and Reis (2006b) model, provided that the solution is not prematurely truncated. The numerical results in section 3.5 below use the undetermined coefficients solution as in equation 3.25.

3.3 Loss Functions

The loss to inattentive agents due to their lack of up-to-date information can be expressed as a combination of the conditional mean square error given their information set and the assumed cost of their updating probability. Further, agents are considered to be making a decision in period t on the basis of the expected loss in the future, as in chapter 2, such that the loss can be expressed as:

$$\begin{aligned} Loss = \quad & \xi\left(1-b\right)E_t\left[\left(d^*_{t+1}-E_t\left[d^*_{t+1}\right]\right)^2\right] \\ & +\sum_{q=2}^{Q}\xi^q\Bigg(b\sum_{m=1}^{q-1}(1-b)^{q-m} \\ & \cdot E_t\left[(d^*_{t+q}-E_{t+m}\left[d^*_{t+q}\right])^2\right]+ \\ & (1-b)^q E_t\left[(d^*_{t+q}-E_t\left[d^*_{t+q}\right])^2\right]\Bigg)+\sum_{q=1}^{Q}\xi^q C_b b^2 \end{aligned} \tag{3.26}$$

for $(b,d)\in\{(\lambda_f,p),(\omega_h,c),(\omega_h,w)\}$ and $*$ denoting a variable with the fully-informed optimal value. Note that for a household the total loss is the sum of this expression over $d=c$ and $d=w$ with $b=\omega_h$, as they wish to optimally choose consumption and the nominal wage, whereas for a firm the loss is this expression evaluated for $d=p$ with $b=\lambda_f$ as it wishes to optimally set price. Additionally, in equation 3.26 λ and ω are

[9] p991, Meyer-Gohde (2010)

subscripted because these are the probabilities chosen by individual firms or households, respectively, rather than the 'aggregate' probabilities held by all other agents and which are taken as invariant to the individual choices. C_b is a term representing the cost of information and this is taken to be specific to each type of agent. Q is the maximum number of periods ahead that agents consider for the loss and the sum indexed by q is therefore the sum of the expected losses in each future period[10]. Within this sum the sum indexed by m is the probability-weighted sum over the possible ages of information that an agent could have in each future period. Considering equation 3.26 for a firm, it seems reasonable that the weighted mean square error must be an acceptable approximation to the relative loss of profit arising from any pricing decisions based on out-of-date information, with the cost term justified on the same grounds as in chapter 2. In the case of households, equation 3.26 assigns equal weight to errors in the choice of consumption and in the choice of labour supply, but this is not necessarily what the household utility function would imply. It might be preferable to determine an appropriate approximation to the household utility with differential weights on consumption and labour supply, into which the expectational error terms could be substituted. This is left for future work (it is not clear whether representing the potential loss to households in this manner would significantly affect the computing time involved in finding the solution).

Using the equations above and the Mankiw and Reis (2006b) solution method it is possible to solve for the expectational error terms relating to the individual decision variables (firm price, household consumption and wage) to get the following expressions[11,12], where the set of shocks is $S = \{\varepsilon, \Delta a, g, \gamma, \nu\}$:

[10]In principle $Q = \infty$, but for computation it is necessary to impose a limit. Note that this is not the same as the limit on the number of lagged expectations included in the model solution. These limits are discussed more in section 3.4.

[11]Recall that $\hat{z}_n(s)$ denotes a response coefficient from the general solution, equation 3.25.

[12]Equations 5.99, 5.100 and 5.101 are used to derive these expressions, in which * is used to emphasize optimality, equivalent to the value under full information, i.e. a superscript of 0.

Firms

$$\tilde{p}^*_{t+q} - E_{t+m}\Big[\tilde{p}^*_{t+q}\Big] = \frac{1}{\beta + \bar{\nu}(1-\beta)} \sum_{s\in S} \sum_{n=0}^{q-m-1} \Bigg(\beta \hat{w}_n(s) + (1-\beta)\bar{\nu}\hat{p}_n(s) + (1-\beta)\hat{y}_n(s)\Bigg) e^s_{t+q-n} + \frac{1}{\beta + \bar{\nu}(1-\beta)} \sum_{n=0}^{q-m-1} \Bigg(-\frac{\beta}{\bar{\nu}-1}\rho^n_\nu e^\nu_{t+q-n} - \sum_{z=0}^{n} \rho^z_{\Delta a}\, e^{\Delta a}_{t+q-n} \Bigg) \tag{3.27}$$

Consumers

$$c^*_{t+q} - E_{t+m}\Big[c^*_{t+q}\Big] = \sum_{n=0}^{q-m-1} \left(\sum_{z=0}^{n+T-q} \rho^z_{\Delta a} + \frac{\hat{y}_n(\Delta a)}{\Omega_n} \right) e^{\Delta a}_{t+q-n} + \sum_{n=0}^{q-m-1} \left(\frac{\psi\beta\rho_g^{n+T-q}}{\psi+1} + \frac{\hat{y}_n(g) - \rho^n_g}{\Omega_n} \right) e^g_{t+q-n} + \sum_{n=0}^{q-m-1} \frac{\hat{y}_n(\varepsilon)}{\Omega_n}\, e^\varepsilon_{t+q-n} + \sum_{n=0}^{q-m-1} \left(\frac{\psi\beta\rho_\gamma^{n+T-q}}{(\psi+1)(\bar{\gamma}-1)} + \frac{\hat{y}_n(\gamma)}{\Omega_n} \right) e^\gamma_{t+q-n} + \sum_{n=0}^{q-m-1} \left(\frac{\psi\beta\rho_\nu^{n+T-q}}{(\psi+1)(\bar{\nu}-1)} + \frac{\hat{y}_n(\nu)}{\Omega_n} \right) e^\nu_{t+q-n} \tag{3.28}$$

Workers

$$w^*_{t+q} - E_{t+m}\left[w^*_{t+q}\right] = \sum_{s\in S}\sum_{n=0}^{q-m-1}\left(\left(c^*_{t+q} - E_{t+m}\left[c^*_{t+q}\right]\right)_n(s) + \hat{p}_n(s) + \frac{\hat{l}_n(s)}{\psi}\right)e^s_{t+q-n} - \frac{1}{\bar{\gamma}-1}\sum_{n=0}^{q-m-1}\rho^n_\gamma e^\gamma_{t+q-n} \tag{3.29}$$

where $\left(c^*_{t+q} - E_{t+m}\left[c^*_{t+q}\right]\right)_n(s)$ represents those terms in $c^*_{t+q} - E_{t+m}\left[c^*_{t+q}\right]$ relevant to the response to a disturbance of type s dated $t+q-n$ (as shown in equation 3.28).

The above expressions make use of the following rearrangement of the technology process:

$$\begin{aligned}
\Delta a_{t+q} &= \rho_{\Delta a}\Delta a_{t+q-1} + e^{\Delta a}_{t+q} \\
(1-(1+\rho_{\Delta a})L + \rho_{\Delta a}L^2) &= e^{\Delta a}_{t+q} \\
(1-L)(1-\rho_{\Delta a}L)a_{t+q} &= e^{\Delta a}_{t+q} \\
(1-L)a_{t+q} &= \sum_{z=0}^{\infty}\rho^z_{\Delta a}e^{\Delta a}_{t+q-z} \\
a_{t+q} - E_{t+q-k}\left[a_{t+q}\right] &= \sum_{b=0}^{k-1}\sum_{z=b}^{k-1}\rho^z_{\Delta a}e^{\Delta a}_{t+q-b}
\end{aligned} \tag{3.30}$$

Given a value of Q, equations 3.27, 3.28 and 3.29 can be used as necessary in the appropriate version of equation 3.26 to evaluate the loss due to potentially out-of-date information. The optimal updating probabilities can then be found as the parameter values which minimize these losses.

The model of Drager (2011) is basically the same as that described by equations 3.2 to 3.12 (and the related linearized equations in appendix A.2.2). However, it differs from the overall model here in the following ways: the labour market is not explicitly considered; monetary policy follows a rule which depends upon the lagged interest rate as well as inflation and the output gap; the updating probability is the same across households and firms but may differ according to which variable is being forecast; and, most importantly, agents choose whether or not to update, rather than choosing a probability. Specifically, agents are assumed to evaluate, in each period, the mean squared errors of past forecasts for each possible age of information and then form a weighted average of these to arrive at a total loss value[13]. The probability that an agent will choose to update is then a function of this total loss value and over the whole population this is equivalent to an aggregate updating probability. A possible difficulty with this is that it requires agents to have information dated t - 1 and earlier in order to determine the past losses, but then those agents who do not opt to update make their other decisions based on information from before t - 1. A second, related, aspect of this model is that in order to derive a 'heterogeneous expectations operator' which is a linear combination of the full- and lagged information-based expectations and which has the desirable property that expectations of each agent about other agents' expectations of the expectations of other agents (and so on) are consistent between agents such that the Law of Iterated Expectations applies to the operator[14], it is necessary to assume the axioms given in Branch and McGough (2009). These axioms are acknowledged to be quite restrictive by the authors, particularly the requirement that higher-order expectations are consistent, but for a model such as this this may not be too restrictive, given that expectations are still rational.

An interesting feature of the model of Drager (2011) is that, due to agents re-evaluating past losses in each period, the proportion of agents who choose to update in each period will vary over time. In particular, it is found that there is a strong positive correlation between the updating probability/proportion and the volatility of variables - the comparative static equivalent of this result is considered in section 3.5. It is also found that agents are more likely to update their forecast of output than their forecast of inflation, and the suggested explanation for this is that this is because the central bank reacts more strongly to inflation than output, so agents 'trust' the central bank to be responsible for limiting inflation variability and concentrate on output themselves. The model here does not have differential updating probabilities for expectations

[13]These weights are assumed in addition to the updating proportions.

[14]Fukunaga (2007) analyses the problem of expectations of expectations in a model similar to that of chapter 2 with a signal extraction problem, but it is unclear whether such an approach would be applicable here.

of different variables, so it is not possible to produce a comparable result; however, it is possible to look at the different equilibria which occur as the monetary policy rule coefficients vary - the results from doing so are presented in section 3.5.

3.4 Computation

The expressions in 3.28 and 3.29 feature T (implicitly in the case of the latter); this occurs as a consequence of imposing a period $t + T$ which is the furthest forward that agents consider when considering the long-run natural output[15]. Given that T occurs as an exponent on terms which are less than one in absolute value any omitted quantities will be small, provided that T is quite large. In the computation of the loss functions the value of Q is limited to 50, in order to reduce the processing time. This may alter the results compared to the case of agents seeking to minimize the mean square error over the infinite future. However, given that in equation 3.26 future losses are discounted and that the probability of a very old information set is likely to be small, the exclusion of periods in the distant future may not have a significant effect on the optimal probabilities. This seems to be the case in practice, with increases in Q beyond 50 having an effect on the numerically calculated optima of typically less than 10^{-8}. Note that the value of Q only applies to equation 3.26, it is not a restriction on the general solution of the model, which allows for a much longer maximum information age.

The endogenous updating solution procedure is the same in principle as in chapter 2. The equilibrium computed is the symmetric Nash equilibrium for each type of agent and the (asymmetric) Nash equilibrium across different types of agent. This means that rather than finding a single equilibrium between an individual and the 'aggregate' for a single type of agent it is now necessary to find an overall equilibrium such that the loss is minimized for firms given the choices of other firms and households, and for households given the choices of other households and firms, where the model must be re-solved for each (λ, ω) tuple[16]. This is therefore a multi-objective optimization problem[17]

[15]The long-run natural output features in the decision due to the consumer's optimality condition being solved forward - see equation 5.98 in appendix A.2.2.

[16]The possibility of multiple equilibria akin to those found in the second chapter will be discussed in the results section

[17]Such problems are also known as vector or Pareto optimization problems, as the outcome is often

with two simultaneous objectives, specifically the intra-type Nash equilibria, with the overall Nash equilibrium being imposed by simultaneity. In principle the computational problem ought to be solved as a multi-objective optimization problem in keeping with the theoretical specification, rather than as a single-objective minimization[18], because the algorithm could converge to, for example, a point such that the equilibrium was very accurate for firms but very inaccurate for households, with this point being far from a satisfactory solution overall. In the absence of analytical solutions it is not possible to say whether, in general, a change in one of the control variables which produced a movement towards one of the two objectives would be likely to lead to a movement towards or away from the other objective. However, it is found that in practice the two objectives are not conflicting, in the sense that it is possible to achieve both to within a satisfactory numerical accuracy by minimization of a single objective function[19], (e.g. all errors at most 10^{-4} and, in many cases, several orders of magnitude less). This reduces the computation time both because the problem is simpler and because the minimization can be performed by a standard algorithm.

3.5 Results

The numerical results in this section use either the calibrated or estimated parameter values from Table 1 in Mankiw and Reis (2006b), which are maximum likelihood estimates in the case of the latter. A constant exception to this throughout is that the updating probabilities are determined endogenously as described above, rather than being fixed to estimated values. Here there are also the information cost parameters, which have no counterpart in the exogenous updating model and are essentially therefore free parameters. In order to have a benchmark case with which the other results can be compared, the information costs are set to the values which result in the updating probabilities estimated by Mankiw and Reis (2006b) being (approximately) the equilibrium,

---

a set of candidate solutions such that it is not possible to move closer to one objective without moving farther away from another.

[18]E.g. $min_{\lambda,\omega}(\bar{\lambda} - \lambda_f^*)^2 + (\bar{\omega} - \omega_h^*)^2$, where $^-$ denotes the 'aggregate' probability chosen by agents of a particular type and $*$ denotes the optimal individual updating probability.

[19]$min_{\lambda,\omega}(\bar{\lambda} - \lambda_f^*)^2 + (\bar{\omega} - \omega_h^*)^2$. The exponent is greater than 1 in order to penalize undesirable/unbalanced solutions such as that suggested in the main text.

endogenous probabilities also, given all other parameter values the same. It is debatable whether the calibrated values are the best to use; for example, the Taylor rule coefficients are estimates from Rudebusch (2002), who uses a data sample from 1987 to 1999, but the data sample used by Mankiw and Reis (2006b) is from 1954 - 2006, in which time it is likely that the monetary policy regime has not been constant - indeed it has been suggested by Clarida and Gertler (2000) that there was a significant shift in policy regime in 1979, with the appointment of Paul Volcker to the chairmanship of the USA Federal Reserve Board. Their results also suggest that for the earlier part of this sample, monetary policy could not be characterized as having been conducted in such a way that the Taylor principle was satisfied, although the form of the policy rule they estimate is somewhat different to that used here. However, as the results here concern calculation of the optimal updating probabilities and of macroeconomic behaviour in response to hypothetical parameter variations, rather than estimation, using these calibrated values may not have a significant impact.

The following table lists the parameter values used for the results presented thereafter, except where stated otherwise.

Parameter	Value	Parameter	Value
ν	34.068	ρ_g	0.93764
γ	4.196	σ_g	0.013929
β	2/3	ρ_γ	0.66682
ψ	4	σ_γ	0.18657
θ	1	ρ_ν	0.62975
ϕ_y	0.33	σ_ν	1.8194
ϕ_p	1.24	ρ_ε	0.9179
C_λ	0.00033	σ_ε	0.0116
C_ω	0.4	$\rho_{\Delta a}$	0.3496
		$\sigma_{\Delta a}$	0.0102

Table 3.1: Benchmark Parameter Values

The interactions between the equilibrium updating probability of each type of agent, those of the other types of agents and the exogenous parameters will be considered first, followed by the impulse responses and simulated moments of the macroeconomic variables.

Figure 3.1 shows the best-response function for each of the types of agent given the benchmark parameterization. In each panel the 'aggregate' probability of that type, i.e. the probability for that type which is used to solve the model to determine the value of the loss function, is on the horizontal axis and the optimal choice of the individual agent given this is on the vertical axis. The left-hand panel pertains to firms and the right-hand panel to households. In each case the symmetric Nash equilibrium is shown by the intersection of the best-response function with the dotted 45^o line. The equilibria are mutually consistent, in the sense that the best-response functions for both types correspond to the losses calculated using the same model solution, based on the same values of λ and ω (the household best response function shown corresponds to the lower of the two firm equilibria, with the benchmark equilibrium being $\lambda_f^* = 0.707$ and $\omega_h^* = 0.207$, approximately the same as in the exogenous case of Mankiw and Reis (2006b)).

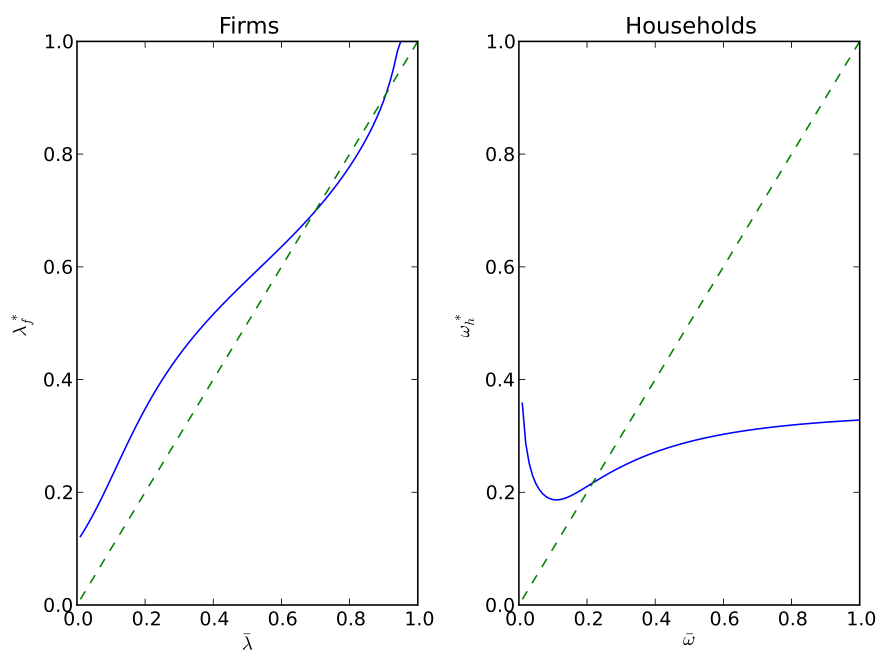

Figure 3.1: Benchmark Equilibrium Updating Probabilities

It can be seen from figure 3.1 that the best-response function of firms is entirely positively sloped, while the best-response function for households is negatively sloped when the updating probability of other households is very low, then once $\bar{\omega}$ increases beyond approximately 0.1 the best-response function is positively sloped, but this slope is less than 45^o. In all of the equilibria computed these properties are exhibited[20]. As discussed in chapter 2, for the model to have multiple equilibria it is necessary that the best-response functions be positively sloped (at least over some range of the 'aggregate' probabilities). In this example the only best-response function which suggests multiple equilibria is that for firms, which has two intersections of the 45^o line. Given that the best-response function for households only intersects once there is only one equilibrium here, and that would also be the case even if the firms' best-response function crossed the 45^o line several times, as each of those intersections would correspond to a different best-response function for households. In general, for there to be multiple equilibria there would have to be more than one value of each of the updating probabilities such that they were an intra-type symmetric equilibrium, allowing for the effect on agents of the other type. In the numerical results produced here there have not been found to be any instances of multiple equilibria, so on the basis of these computed results multiplicity of equilibria does not seem to be a significant issue for this model, whereas it was for the simpler model of chapter 2.

Further possible evidence that the equilibrium will be unique in this model is provided in figure 3.2, which shows the surface of equilibrium firm probabilities in the upper plot and the surface of equilibrium household probabilities in the lower plot, the surfaces being the sets of equilibria as both $\bar{\lambda}$ and $\bar{\omega}$ vary, given the benchmark parameterization. It can be seen that both surfaces vary in all dimensions. For there to be multiple equilibria overall it would have to be the case that there were several points on these surfaces for each type of agent where the value of the optimal updating probability was equal to

[20]However, given the number of parameters in the model, the proportion of the parameter space that it has been feasible to investigate numerically is small. Therefore it cannot be claimed that these signs of the slopes of the best-response functions are generic properties. Similarly, the other results described below ought not to be thought of as generic.

the type aggregate probability and this value was mutually compatible with a point or points on the surface applicable to the other agent type which was also a symmetric equilibrium for that agent type. It is not immediately apparent from figure 3.2 whether a number of such pairs of points may exist, however given that the majority of the surface applicable to households, shown in the lower plot, lies beneath the '45^o plane' there will be a relatively small proportion of the household updating probabilities which are compatible with symmetric equilibrium, so this in turn will restrict the area of the surface applicable to firms (shown in the upper plot) for which equilibria are possible. For other parameter values it may be possible for the equilibrium surfaces to be such that multiple general equilibria are likely, but this has not been observed in any of the results produced.

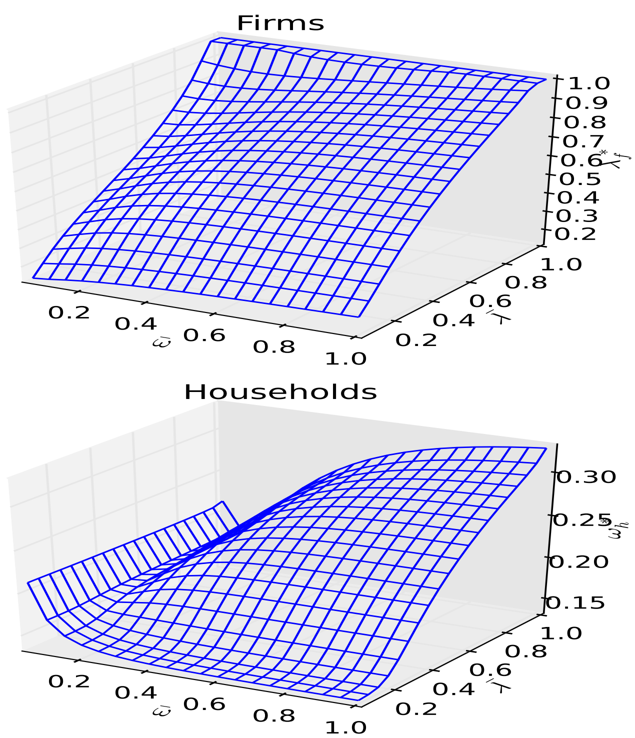

Figure 3.2: Benchmark Equilibrium Probability Surfaces

In order to consider the strategic inter-relationships between different types of agents, figure 3.3 shows the variations in the equilibrium probabilities as the costs of information acquisition and processing vary. In each panel the cost parameter for one type of agent varies, with the other held constant. This therefore shows not only the direct effect of cost on the optimal choice but also, given that each cost parameter features only in the loss function of the relevant agent type, the indirect effects between agents of different types (this relationship is also observable from figure 3.2).

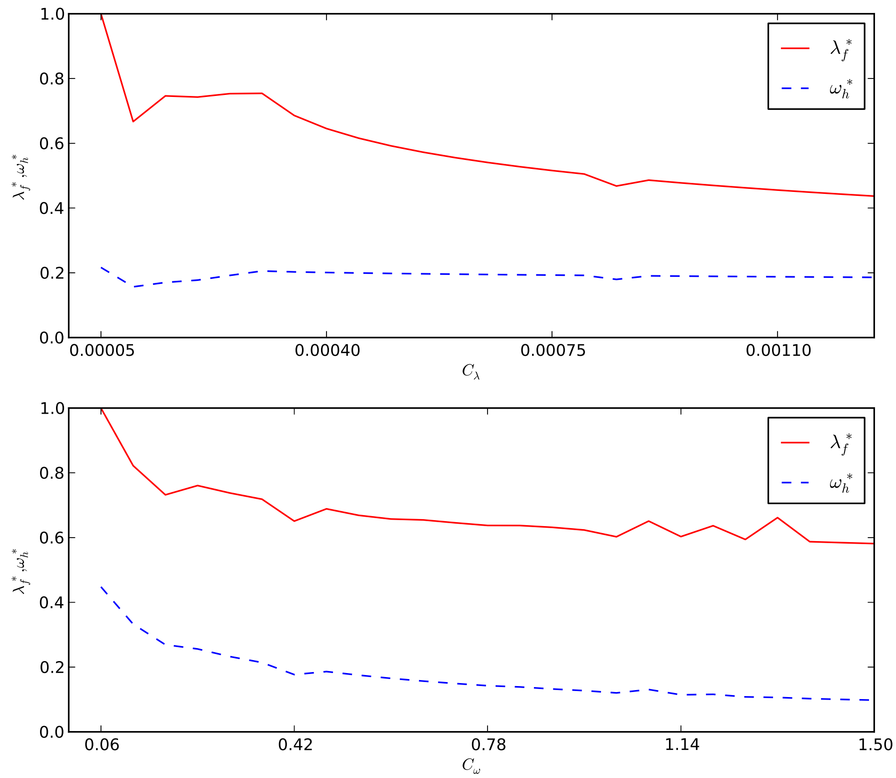

Figure 3.3: Equilibrium as Information Cost Varies

Beginning with the top panel of figure 3.3, it can be seen that as the cost to firms of receiving and processing new information increases, the optimal probability with which firms do so declines. It can also be seen that the optimal choice of households is almost invariant to this change in firms' updating probability. This is perhaps a surprising result, because it would be expected that in a general equilibrium model a change in updating probability by one type would affect the choices of the other type, even though the types of inattentive agent do not directly interact. It appears that, at least under this parameterization, the size of the effect of firms' choice on households is very small. The lower plot shows the equilibria as the information cost to households varies. The optimal probability for households declines as the cost increases, while the optimum for firms declines at a similar rate. An explanation for this may be that as households update less often, on average, the variability of the real wage decreases and so the variability of real marginal cost decreases, allowing firms' pricing departments to update less often without suffering a greater departure from the optimal price than is compensated for by the reduced information costs. It is curious that the reverse of this does not induce a response by workers to changes in the updating probability of pricing departments. Overall, figures 3.1 and 3.3 suggest that with respect to the choice of updating probability the strongest strategic relationships in the model are those between agents of the same type, rather than between agents of different types. In particular, firms seem to have the strongest such relationship.

Considering next the relationship of the optimal updating probabilities with the volatilities of the exogenous shocks, figure 3.4 shows the equilibria resulting from differing standard deviations of the shocks to technology growth, the labour substitution elasticity and the goods substitution elasticity, while figure 3.5 shows the equilibria resulting from differing standard deviations of the shocks to the policy (monetary and fiscal) variables. The first panel of figure 3.4 shows that firms and households respond quite strongly to an increase in the volatility of the shock to technology growth. Within the range of standard deviations used here, it becomes optimal for firms to be fully informed, perhaps because the level of technology, which is permanently affected by shocks to technology growth, features in the optimal price condition. Similarly, for households solving the Euler equation forward reveals a dependence on the long-run natural rate of output, which is permanently affected by such shocks. This plot and several of the other plots showing the varying equilibria feature some quite sharp kinks. These are not likely to be numerical errors, as only those computational results which satisfied the optimization tolerance were retained, so numerical instability should not be a problem. However, these could represent the existence of local minima around the global minimum (assuming that this would smoothly follow the trend implied by the majority of the results). Alternatively these kinks could represent the true outcome, although this seems to be unlikely given the general pattern of the results.

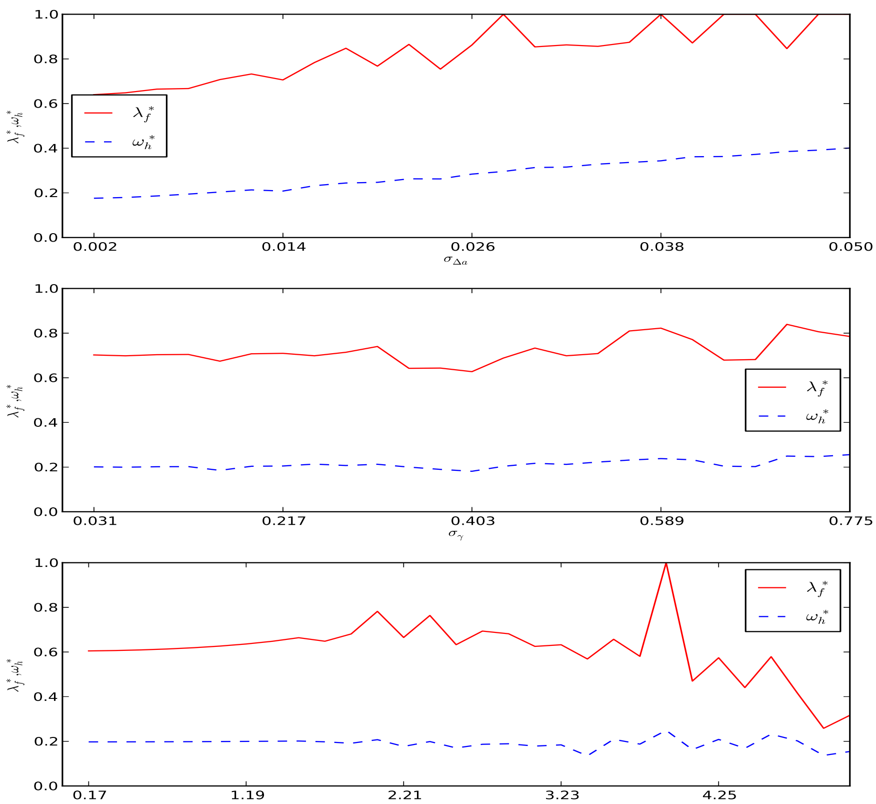

Figure 3.4: Equilibrium as Volatilities of Non-Policy Shocks Vary

The middle panel of figure 3.4 shows that the response of all agents to increases in the standard deviation of the shock to the labour substitution elasticity is relatively small, with some increase in the optima for firms and households, but not of a sufficiently large amount to change the properties of the model significantly. The bottom panel of figure 3.4 shows that households respond relatively weakly, if at all, to increases in the volatility of the goods substitution elasticity and that for firms it may even be optimal to be reduce the updating probability as the volatility increases. This is a curious result which goes against the intuition that a more volatile economy would induce greater attentiveness and that firms who experience strategic complementarity in price-setting would wish to know more about the prices set by rival firms when those prices were more variable.

The upper graph in figure 3.5 shows the effects of changes in the standard deviation of the shock to the monetary policy rule. Despite monetary policy being modelled as exogenously given, rather than the behaviour of an optimizing policy-maker, the volatility of the shock could perhaps be thought of as indicative of the degree of competence/success with which monetary policy is implemented, or the degree to which monetary policy is predictable. Under this interpretation, the graph suggests that less tightly-controlled monetary policy leads private agents to update their information more frequently. This effect is strong for both households and firms, with certain updating being optimal for the latter when the shock is very variable (at some value above twice the benchmark standard deviation). Even within a small neighbourhood of the benchmark value of σ_ε (0.012) an increase produces quite a large increase in the optimal probabilities for households and firms. Indeed, of all of the parameter variations considered in this section, it appears that the strongest dependence of the household updating probability is on the volatility of the monetary policy shock.

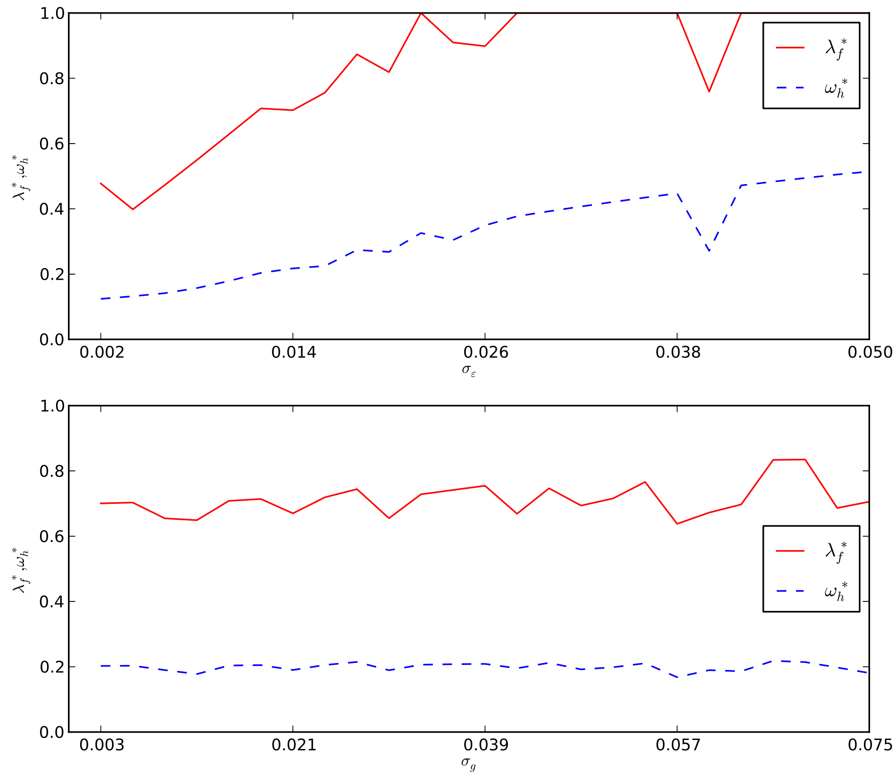

Figure 3.5: Equilibrium as Volatilities of Policy Shocks Vary

The lower graph of figure 3.5 above shows the effects of changes in the standard deviation of the shock to the variable which indirectly represents fiscal policy. It can be seen that firms and households respond very little to increases in this standard deviation (the benchmark value of which is roughly the same as for the monetary policy shock). Taken together, the results in figure 3.5 suggest that if policy was to become more accurately implemented in order to reduce macroeconomic volatility, such a change would be more effective in the case of monetary policy, as this would induce the larger change in average updating frequency on the part of private agents, and this may in turn produce a greater smoothing of the impact of shocks[21].

[21] This possibility is discussed further below.

Figure 3.6 shows the equilibrium optimal updating probabilities for different values of the coefficients on the output gap (in the upper panel) and inflation (in the lower panel) in the monetary policy rule. In both cases an increase in the responsiveness of the nominal interest rate leads to decreases in the updating probabilities of firms and households of similar magnitudes, although the updating probability of households does not change by a large amount. These results accord with most of those for the variability of shocks, in that in a more stable economy (corresponding to larger policy coefficients) private agents do not face as large expectational mean square errors resulting from out-of-date information and so can afford to choose lower updating probabilities. However, unlike the case of changes to the standard deviation of exogenous monetary shocks, in this case the choice of households appears to be almost invariant with respect to the policy rule coefficients, whereas it varied significantly with respect to the shock volatility. Firms respond strongly to variation in either monetary policy rule coefficient.

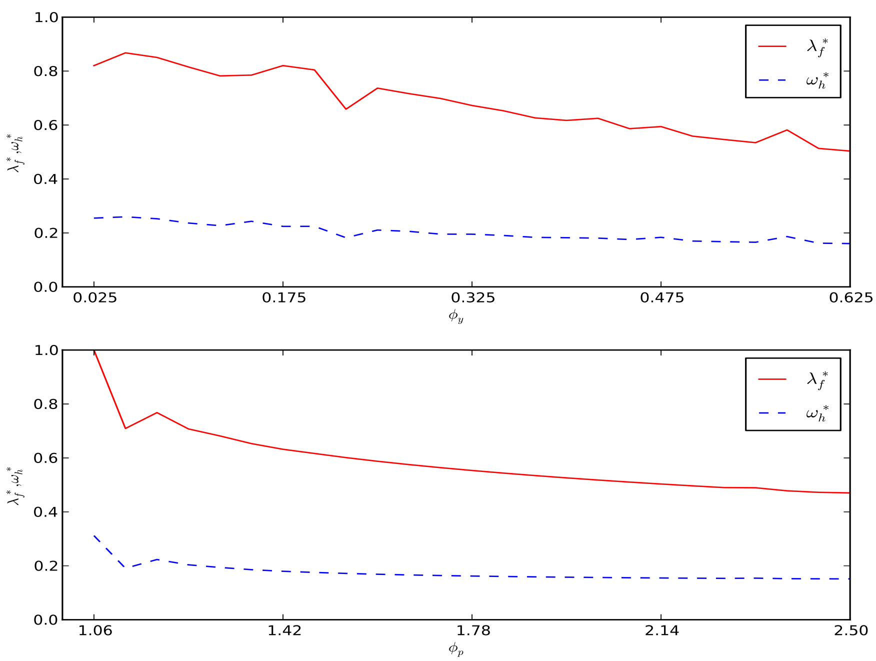

Figure 3.6: Equilibrium as Taylor-Rule Coefficients Vary

Figure 3.7 shows the impulse responses of output and inflation for a shock to monetary policy, firstly, and secondly for a shock to technology growth, given the benchmark parameterization. As a positive shock to monetary policy is taken to reduce the nominal interest rate, the responses are positive for both output and inflation, with the maximal inflation responses occurring three to five quarters after the impact date and there continuing to be some response for a long time after the initial impact. Had all agents been perfectly informed the maximal responses of inflation would be expected to occur immediately, with a rapid decline in the responses over time, as in the standard real business cycle model without any nominal rigidities.

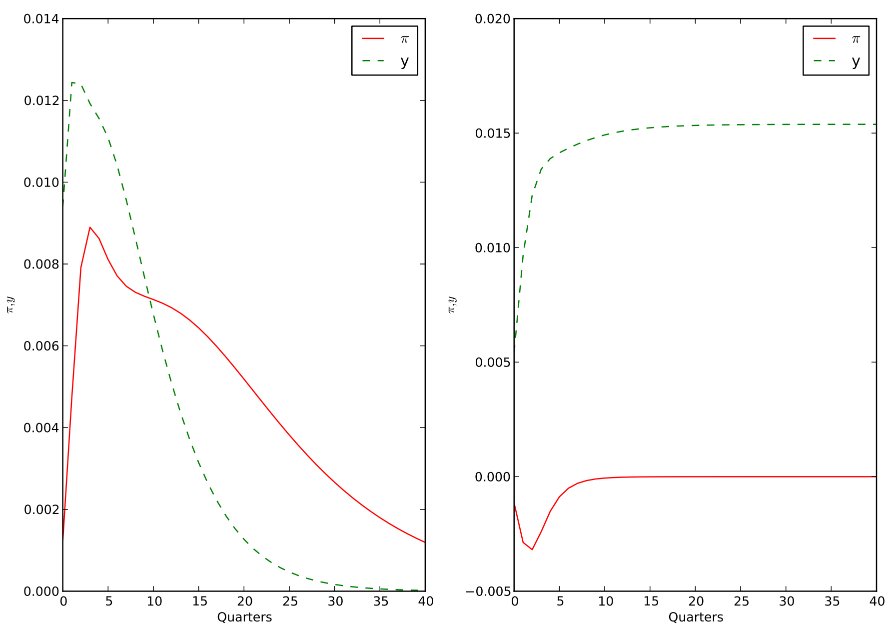

Figure 3.7: Benchmark Impulse Responses to Monetary Policy and Technology Shocks

In the case of a shock to technology growth (shown in the right-hand plot), it can be seen that the maximal response of inflation occurs with a slight delay and then the effect soon dies out. In contrast, the response of output shows a permanent increase as a result of a positive shock to technology growth, because the level of technology, upon which output depends, is permanently increased thereby.

To investigate further the macroeconomic effects of changes in either the coefficients of the monetary policy rule or the variability of the shocks to the policy rule, the following considers the simulated standard deviations of output and inflation and the impulse responses for output and inflation, given changes to monetary policy. It was asserted above that the effects of such policy changes on the equilibrium updating probabilities would have macroeconomic consequences which distinguish the endogenous sticky information model from that with exogenous sticky information. In order to demonstrate this, Table 3.2 shows the simulated standard deviations of output and inflation for an increase in the policy rule coefficient on inflation from 1.24 to 1.7 or 2.3. In each case the same series of realized shocks is used and the model is simulated for 210 periods $10,000$ times, with the average standard deviation being the result shown; results using other draws of the shocks indicate that this is a sufficiently large number of simulations to ensure that the results shown in Table 3.2 are both qualitatively and quantitatively representative of the outcome of this policy change. The numbers shown in brackets in Table 3.2 are the percentage changes in the standard deviations relative to the benchmark case. For each of the increased policy coefficients the table shows the outcome under exogenous sticky information, i.e. if λ and ω did not change, and the outcome under endogenous sticky information.

	σ_y	σ_π
$\phi_p = 1.24$		
Exogenous SI $\lambda = 0.707$, $\omega = 0.207$	0.127	0.366
$\phi_p = 1.7$		
Exogenous SI	0.1241 (-2.3)	0.113 (-69.1)
Endogenous SI $\lambda = 0.568$, $\omega = 0.165$	0.1244 (-2.0)	0.090 (-75.4)
$\phi_p = 2.3$		
Exogenous SI	0.1234 (-2.8)	0.074 (-80)
Endogenous SI $\lambda = 0.486$, $\omega = 0.153$	0.1236 (-2.7)	0.056 (-85)

Table 3.2: Simulated σ_y and σ_π for Varying ϕ_p

It can be seen that in the exogenous updating probabilities case the first policy change reduces the variability of both output and inflation, by 2.3% and 69.1% respectively. In the case of endogenous probabilities, the reductions are 2% and 75.4%, with quite large changes in the probabilities chosen by firms and households. For the second (greater) policy change considered, output standard deviation decreases by 2.8% and inflation variability decreases by 80% in the exogenous sticky information model and in the endogenous sticky information model output variability decreases by 2.7% and 85%. These results suggest that the reduced attentiveness of price- and wage-setters following the increase in policy responsiveness to inflation significantly augments the effectiveness of the policy change with respect to inflation volatility, but at the 'cost' of a diminution of the reduction in the variability of output[22]. The results for the two policy changes considered also suggest that this difference between the exogenous and endogenous updating models may be greater for smaller changes in the policy coefficient.

[22]As specified above, the updating probability is chosen as a once only decision, after which it is fixed rather than being continuously re-optimized. However, for the purpose of this 'policy experiment' it has been supposed that a sufficiently large change in policy would exceptionally justify all agents becoming aware of the change immediately and choosing a new updating probability. Similar justification is appealed to in the case of a change in the variance of the monetary policy shock, as considered below.

Figure 3.8 shows the impulse responses of output and inflation to a monetary policy shock after the policy changes simulated above ($\phi_p = 1.7$ and $\phi_p = 2.3$), allowing for the changes in the optimal updating probabilities following the change. It can be seen[23] that the effects of re-optimized probabilities on the response of output are that the peak response value is lower and this maximum occurs one period later than under the benchmark probabilities. That the largest single-period response of output is lower than in the benchmark case may suggest that the reduction in inflation variability achieved by a more responsive monetary policy does not come at the expense of a greater variability of real variables[24]. The effects of endogenous re-optimization on inflation are that the greatest response is smaller and occurs one or two periods later, relative to the unchanged probabilities case.

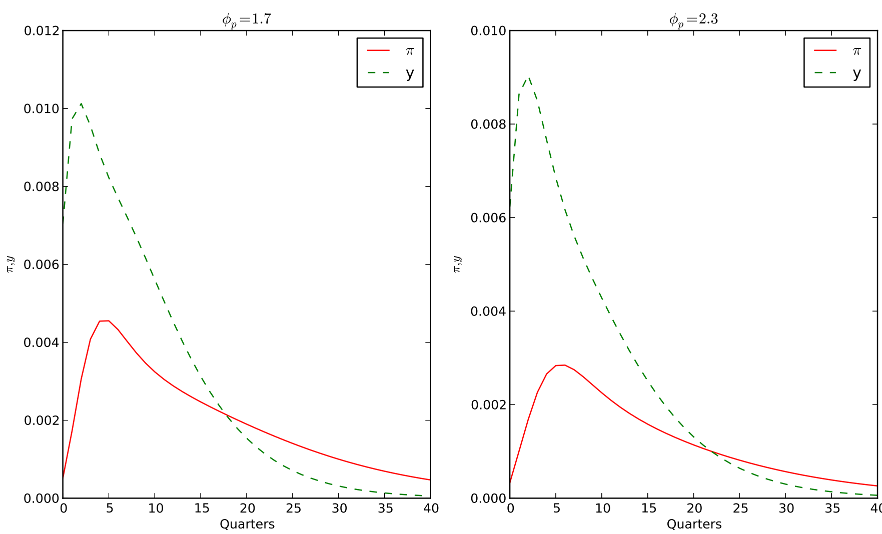

Figure 3.8: Impulse Responses under Stricter Monetary Policy

These results support the findings of altered standard deviations in Table 3.2[25] and show how the dynamics of the responses are affected, with the decreases in attentiveness

[23] By comparison with the left-hand graph in figure 3.7

[24] Although the comparison of the exogenous and endogenous sticky information models in table 3.2 suggests that there is a secondary trade-off between nominal and real rigidity.

[25] And also imply that the results are unlikely to depend significantly on the specific series of shocks used in the simulation.

	σ_y	σ_π
$\sigma_\varepsilon = 0.012$		
Exogenous SI $\lambda = 0.707$, $\omega = 0.207$	0.127	0.366
$\sigma_\varepsilon = 0.003$		
Exogenous SI	0.122 (-3.9)	0.058 (-84)
Endogenous SI $\lambda = 0.517$, $\omega = 0.131$	0.121 (-4.7)	0.030 (-91.8)
$\sigma_\varepsilon = 0.03$		
Exogenous SI	0.150 (+18)	0.533 (+50)
Endogenous SI $\lambda = 1.0$, $\omega = 0.393$	0.138 (+8.7)	0.90 (+146)

Table 3.3: Simulated σ_y and σ_π for Varying σ_ε

resulting in increased delays of the maximal effects.

The discussion of figure 3.5 suggested that a change in σ_ε (the standard deviation of the shock to the monetary policy rule) could be interpreted as a change in the extent to which monetary policy is successfully implemented. In order to consider the effect of endogenous updating coupled with such a change upon the behaviour of macroeconomic variables, Table 3.3 shows the simulated standard deviations of output and inflation, initially for the benchmark case, then for both fixed and re-optimized probabilities given two alternative values of σ_ε. Given that it is σ_ε which is changing, the simulation uses two different series of shocks, corresponding to the two different values of σ_ε used. However, these two series are both based upon the same series of draws from the standard normal distribution. Again, 10,000 simulations of 210 periods are used and the results suggest that this is enough to ensure that the results are not strongly influenced by the particular draw of shocks used. The comparison between exogenous and endogenous updating in each of the the altered volatility cases is based on common realizations of the shock. The numbers shown in brackets in Table 3.3 are the percentage changes in the standard deviations relative to the benchmark case.

It can be seen from the upper panel of Table 3.3 that, as would be expected, a reduction in the volatility of the shock, holding the values of the updating parameters constant, directly reduces the volatility of both output and inflation (by 3.9% and 84%, respectively). Then in the case of private agents re-optimizing their updating probabilities in response to the change in variance, it can be seen that all agents choose lower probabilities (0.517 for firms and 0.131 for households). The effects of this re-optimization are that the decline in the volatility of output is slightly more than in the exogenous case[26] and the decline in the volatility of inflation is also larger than in the exogenous case. An implication of these effects is that an attempt to more closely follow the policy rule, perhaps with the intention to achieve a particular reduction in inflation volatility could, if based on the assumption of fixed inattentiveness, be greater than was necessary to achieve the target. The lower part of Table 3.3 shows the outcomes for an increase in the volatility of the shock to monetary policy. It can be seen that in this case the optimal probability for firms would be 1, i.e. full information in every period, while for households the optimal probability would almost double, to 0.393. Compared with the exogenous information case, the standard deviation of output would increase less while the standard deviation of inflation would increase much more. Overall this effect of an increase in monetary policy variability on output and inflation volatility is quite different in the endogenous sticky information model.

The responses of output and inflation to monetary policy shocks given the changed volatility thereof are shown in figure 3.9, with the responses for a reduced volatility shown in the left panel and those for an increased volatility shown in the right panel. It can be seen that the greatest response of output occurs after three periods in the first case and after one period in the second. This is in line with the changes given different updating probabilities seen in figure 3.8. In both cases the greatest response of inflation occurs after five periods, but in the case of increased σ_ε the initial response is much closer to the peak response than has been the case previously. It is notable that even when firms are fully informed there is some delay in the response of inflation, presumably because the behaviour of households with outdated information imposes some lag on the

[26]This is contrary to the results in table 3.2, where the endogenous case decline in output volatility was lower than in the exogenous case. A possible explanation of this is that here the induced change in the household updating probability is greater and this might be expected to reduce output volatility, although it is unclear why this should effect should not have also influenced the results in table 3.2.

transmission to consumption and wages, which then feed back into the price level over time.

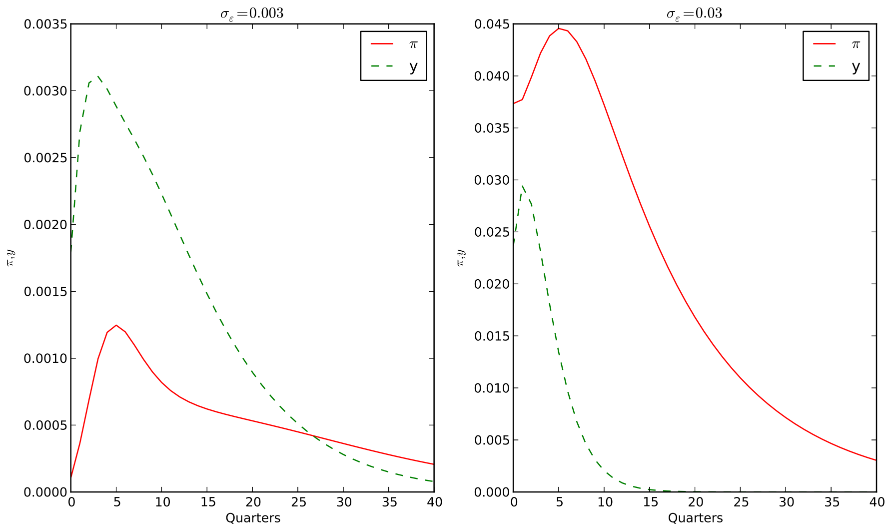

Figure 3.9: Impulse Responses for Different Monetary Policy Error Variability

In comparison with the benchmark impulse responses shown in figure 3.7 less frequent average updating appears to transfer the effects of the shock from inflation to output and vice versa (although to some extent any such effect may be conflated with the change induced by the initial parameter change). This might be expected, in that if agents had complete information prices and wages would incorporate shocks immediately and conversely if new information was never received then prices and wages would not adjust at all to shocks, so shocks would act entirely through output. This argument may be complicated by the inclusion of an inattentive agent (the consumer) who chooses quantities rather than prices, but in general the behaviour of the model shown in figure 3.9 (and to a lesser extent figure 3.8) seems to support it.

3.6 Conclusion

This chapter has shown how a quite detailed DSGE model featuring sticky information can be extended to include the endogenous determination of the probabilities with which agents receive new information. The consequences of this extension for the results of the model have been shown to include a dependence of the optimal updating parameters upon the variances of the exogenous shocks, the coefficients of the rule representing monetary policy and the choices made by other agents (with the strategic effects being strongest between agents of the same type). Furthermore, the changes in these probabilities have been shown to produce significant changes in the dynamic response of the economy to disturbances. In addition, endogenous updating has been shown to have implications for the effectiveness of changes in monetary policy, with a failure to consider such effects potentially leading to 'over-shooting' of policies intended to reduce inflation volatility or limited success when attempting to reduce output volatility. All of these features are absent from the models of exogenous sticky information, so this may contribute to the understanding of some actual changes in the macroeconomy. Ideally the model could be extended to feature continual adjustment of updating, reflecting policy developments over time, rather than the extraordinary re-optimization imposed above. However, the results in this chapter are still likely to be indicative of the ways in which the behaviour of the economy would respond to policy changes. In particular, the results that greater variability of shocks (and hence of macroeconomic variables) induces more frequent updating on average, and that a monetary policy which is more responsive to inflation induces less frequent updating on average, are supported by the results in Drager (2011), which does allow dynamic re-adjustment of updating probabilities[27].

[27] As discussed in section 3.3, the updating probabilities in that model are not exactly equivalent to those considered here.

Chapter 4

Estimation

4.1 Introduction

The previous chapter considered the theoretical behaviour of the model economy with sticky information, given a set of parameter values based on some of the estimation results of Mankiw and Reis (2006b), which were for exactly the same model with the exception of an exogenous updating probability and potentially different updating probabilities between consumers and workers. This chapter now attempts to estimate the model allowing for endogenous sticky information. A range of different estimation methods have been used in the literature to assess the ability of models featuring sticky information to fit macroeconomic data and to compare such models with New Keynesian DSGE models featuring sticky prices, typically with Calvo price updating. The results of these studies are quite varied and are discussed in more detail in section 4.2 below; in general the evidence does not conclusively favour one model over the other, nor does it fully support sticky information as a model feature, as far as data-matching ability is concerned.

The methods used here to estimate and evaluate the sticky information model are simulated method of moments, structural vector autoregressions with minimum distance estimation and indirect estimation. The particular application of these methods is described in section 4.3, as is the data sample used, and the results are then presented and discussed in section 4.4. Section 4.5 concludes this chapter by considering the overall empirical performance of the sticky information model which was presented and analysed in chapter 3.

4.2 Discussion of Literature

Mankiw and Reis (2006a) were, to the author's knowledge, the first to estimate the parameters of a model featuring more than one type of sticky information-constrained agent, the particular model in question being a flexible price model with the addition of potentially outdated expectations, due to exogenous updating probabilities that are variously applicable to the three types of inattentive agent. The model has only two types of disturbance, a shock to technology and a shock to monetary policy. The model was estimated using the simulated method of moments[1] and quarterly data for the USA non-farm business sector spanning 1954 Q3 to 2005 Q3. The particular moments of the data which were used for the estimation were 'three key facts'[2]: the correlation between annual inflation and the mid-year level of output (inflation procyclicality); the ratio of the standard deviation of hourly real wages to the standard deviation of real output per hour (smooth real wages) and the ratio of the standard deviation of quarterly output growth to (half) the standard deviation of annual output growth (gradual response of output to shocks). These three moments were chosen because it was argued that these cannot be matched satisfactorily by any parameterization of the model without sticky information. Interestingly, the results are that in order for the predicted model moments to match these three empirical moments it is necessary that firms, consumers and workers act subject to sticky information; a model with information constraints on only two or fewer of the three agent types does not fit all three of these moments of the data. The results for the updating probabilities are $\lambda = 0.52$, $\omega = 0.66$ and $\delta = 0.36$[3]. As will be discussed, these estimates for worker and consumer updating are higher than is often

[1]The weighting matrix used was the identity matrix, hence each moment was given equal weight in the optimization objective.

[2]See Mankiw and Reis (2006a) pages 164-165.

[3]In their model the parameter δ is the probability of a consumer receiving new information/updating the consumption plan. Here there is only a single updating probability for households, which is equivalent to workers and consumers having the same updating probability.

found elsewhere. Mankiw and Reis (2006a) note that if they impose the restriction of all agents having a common updating probability then the estimate is 0.57 and the model is less well able to fit the empirical moments than in the case of type-varying probabilities, but still better able to do so than the model without information frictions.

Following their development of the model used in chapter 3, Mankiw and Reis (2006b) proceed to estimate the model parameters, without conducting any comparison to other models. Given that the model features five different disturbances and that their solution method solves the model in terms of five variables[4], the method employed is maximum likelihood estimation (initially), followed by the imposition of prior distributions for the parameters and Bayesian estimation. Again the data sample used is quarterly for the USA non-farm business sector, from 1954 Q3 to 2006 Q1. The values of the technology shock persistence and volatility were estimated separately, given a calibrated value of the labour share (β) and the values of the monetary policy shock persistence and volatility were estimated separately given calibrated values of the coefficients of the monetary policy rule. As noted above in chapter 3, these calibrated coefficient values were taken from Rudebusch (2002) but may not have been an entirely satisfactory choice, as the data period for which they were estimated was significantly shorter than the data sample used by Mankiw and Reis (2006b). Given this and the other debatable aspects of monetary policy rules (such as whether a lagged interest rate term should be included or whether it is reasonable to allow policy to respond to current values[5]), the estimation results in section 4.4 allow for the policy coefficients (and consequently the parameters of the monetary policy shock process) to be estimated rather than pre-specified, while maintaining the same form of policy rule as Mankiw and Reis (2006b). The estimated updating probabilities for this paper were $\lambda = 0.702$, $\omega = 0.195$ and $\delta = 0.184$. Compared to the results from the moment-matching estimation, the average frequency of updating by firms is higher, while the average frequency of updating by workers and consumers is much lower (in the case of workers, this estimate implies that, on average, they wait an extra year between information updates).

Coibion and Gorodnichenko (2011) develop a model in which each price-setting firm

[4]See equation 3.25 above.

[5]Taylor (1999) considers these and other issues relating to the specification of monetary policy rules.

may be subject to one of several different types of pricing rigidity; prices may be set in accordance with Calvo sticky prices, sticky information, backward-lookingness (rule-of-thumb indexation) or fully-informed and flexible price-setting. Apart from this possibility the model is quite similar to that considered in chapter 3, except that the monetary policy rule allows for interest rate smoothing and consumption is habitual. A further significant difference is that here wages are fully flexible, unlike the model of the previous chapter[6]. In order to reduce the number of lagged expectation terms the restriction of a maximum information age of twelve quarters is placed upon those firms with sticky information. However, the estimates do not change significantly when this maximum is altered to twenty-four quarters, suggesting that the truncation is not biasing the results, even though it might be thought that limiting the maximum information age but not the maximum time since a price reset (for those firms subject to sticky prices) might bias the results in favour of sticky prices as a source of persistence. The variables used for estimation are the inflation rate, output growth and the nominal interest rate, and the estimation method used is matching moments; specifically, the moments targetted are the first four autocovariances of each of the data series. The method is justified on the grounds that the autocovariances are economically meaningful quantities, such that the results can be interpreted in terms of which dimensions the model fits well or poorly, and that the comparative exercise of Ruge-Murcia (2007) finds that this method performs well against alternatives such as maximum likelihood. The data used is for the USA from 1984 - 2008, with this date range chosen because it is considered to be a period in which a single policy regime/structure existed. As this model allows for different types of firm, it is possible to estimate the proportion of firms of each type within the economy. The base estimates are that 62% of firms have sticky prices, 21% sticky information, 8% flexible prices and 9% set prices by indexation. The estimate of the price reset probability is 0.19 and the estimate of the information update probability is 0.05, which was the lower bound imposed upon the estimation. To the extent that this estimate seems to be inconsistently low in comparison to all of the other studies considered, this raises the possibility that their estimation does not reflect the 'true' probability, or that there

[6] The introduction of rigidity to wage-setting is highlighted as a desirable extension by Coibion and Gorodnichenko (2011) in their conclusion.

is some mis-specification in the model. This caveat aside, these results seem to suggest that sticky prices are a more common/important feature than sticky information. A comparison of the one-year forecast error variance decompositions for models which are restricted to a single type of price-setter shows that for a pure sticky information model a greater proportion of variance is attributable to monetary policy shocks and a lesser proportion of variance is attributable to technology shocks, relative to the pure sticky price model. The explanation offered for this is that following a contractionary monetary policy shock the response of the sticky price or hybrid models is for both inflation and output growth to decline strongly immediately then to have a diminshing response over subsequent periods, whereas for the sticky information model the response of inflation to such a shock is 'hump-shaped' but the response of output growth is not, leading to a correlation between these two variables which is contrary to the data. Two objections to this explanation could be as follows: firstly, the model in Coibion and Gorodnichenko (2011) only applies sticky information to price setters, but Mankiw and Reis (2006a) showed that when the model allows for sticky information across all agents the procyclicality of inflation can be matched by the model; secondly, corresponding to the unusually low estimate of the updating probability, the inflation impulse response function with respect to monetary policy shocks appears to achieve a maximal response approximately twelve quarters after the occurrence of a shock, which is a considerably later peak than is typically found. Therefore, these results are perhaps less directly comparable to the results of section 4.4 than some of the other results discussed in this section. A final interesting result from that study is that the ability of the model to satisfactorily match the covariances of the data may in part depend upon strategic complementarity between all price-setters (including those of different rigidity type), as the pure sticky information model produces a lower degree of persistence than the general model (although as previously noted, this may be due to the lack of 'pervasiveness' of sticky information in this model).

Dupor et al. (2010) construct an analogue to the New Keynesian Phillips Curve which features 'dual stickiness', i.e. firms have both a probability of being able to reset their price in each period (as with Calvo sticky prices) and a separate probability of receiving new information and updating their expectations in each period[7]. An interesting feature of this specification is that the resulting inflation equation automatically includes a lagged inflation term, whereas for the 'hybrid' New Keynesian Phillips Curve lagged inflation is usually included by imposing the restriction that those firms which are not able to reset to the optimal price instead adjust their prices according to lagged inflation. The econometric method used to compare these two models of inflation is minimum distance estimation - specifically, a VAR of inflation, real marginal cost and the output gap is estimated, then these estimates are used in the models to produce simulated series of inflation. Minimizing the distance between actual and predicted inflation then provides estimates of the stickiness parameters. As this minimization uses simulated inflation Dupor et al. (2010) do not use the asymptotic distribution of the estimator, but instead bootstrap the distribution by generating 10,000 inflation series and estimating the stickiness parameters for each series. Using quarterly USA data from 1960 to 2007, they find that both the information arrival probability and the price reset probability are significantly different from zero. When both are present the point estimates are SP = 0.14 and SI = 0.42 (for SP the sticky price probability and SI the sticky information probability). These may suggest that persistence in macroeconomic variables is due more to sticky prices than sticky information. Indeed, when the restriction SP = 0 is imposed, the estimate of SI is 0.1, suggesting a possible downward bias of the updating probability when sticky prices are omitted. However, it is also found that the closeness of fit of predicted inflation to actual inflation is best when both frictions are included, so it would not be correct to conclude that sticky information is merely a proxy for sticky prices. Dupor et al. (2010) also conduct their estimation for sub-samples of the date, firstly for the range 1960 - 1979 and then for 1984 - 2007. The estimates are not qualitatively different from the whole-sample results and the differences in the estimates between the two sub-samples are quite small. Finally, the model is re-estimated to allow for strategic complementarity in price-setting and it is found that this increases the estimate of SI from 0.42 to 0.7, which accords with the theoretical results in chapter 2, where it was shown that in that model there is a positive relationship between the

[7] In contrast to the specification of Coibion and Gorodnichenko (2011), where each firm was only subject to one type of rigidity.

degree of strategic complementarity and the optimal updating probability[8].

(Knotek, 2010) differs from the two papers previously discussed in that he develops a model of price-setting firms under sticky prices and sticky information in which pricing is state-dependent, rather than time-dependent. The estimation approach is also somewhat different from those used above. The estimation is conducted using the simulated method of moments and the empirical moments to be targetted are not only macroeconomic but also microeconomic; the macroeconomic moments are the coefficients of an estimated regression of inflation on four lags of inflation and the output gap, while the microeconomic moments are the mean duration between price changes and the mean size of price changes. Both sets of empirical moments are derived from USA data covering 1983Q1 to 2005 Q4. These sample dates are chosen because of the availability of the relevant microeconomic price data and because the period is considered to be structurally stable. For reasons of tractability, an upper limit of eight is imposed upon the age of information. The estimated updating probability is 0.297, which is quite close to the value used by (Mankiw and Reis, 2002) in a 'Phillips curve only' model, but lower than the values estimated in the other studies featuring multiple inattentive agents, as discussed. Comparisons between models featuring only one of the two frictions suggest that the combined model is best able to match the selected moments. A possible limitation of these results is that the information distribution is quite severely truncated. (Meyer-Gohde, 2010) demonstrated the potential distortion of the numerical properties of a sticky information model when the maximum age of information is restricted and this ought to be included in any consideration of these estimation results, as well as the results of the other studies where a relatively small maximum age of information is imposed.

Carrillo (2012) considers two model variants: one in which prices and wages are set subject to Calvo resetting probabilities and those prices and wages which are not reset in a given period are indexed to past inflation and consumption is habitual; and a second in which the only form of rigidity is that prices, wages and consumption are set subject to sticky information. This latter model is therefore essentially the same as that used in chapter 3 and section 4.4 below, with the exception of the exogeneity of the

[8] This finding also accords with those of Coibion and Gorodnichenko (2011), discussed above.

updating probabilities. The econometric approach employed here has two stages - firstly a structural vector autoregression (SVAR) is used to identify the structural monetary policy and technology shocks and the empirical impulse responses to these shocks, then the model parameters are estimated by minimizing the distance between the empirical impulse responses and the theoretical impulse responses produced by the models. The two structural shocks are identified separately using different SVARs and quarterly data for the USA non-farm business sector from 1954 - 2007. The monetary policy shock is identified by imposing short-run zero restrictions - specifically that the shock must have a zero contemporaneous effect on output, inflation, and nominal and real wage growth. This restriction that the policy shock should not affect 'current period' variables is fairly common, but is not consistent with the model of chapter 3. Carrillo (2012) restricts the information set in the model such that 'current' information at period t contains the nominal interest rate of period t - 1 rather than t, so that the model is consistent with the identifying restriction used in the estimation. The technology shock is identified using the long-run restrictions[9] that only the technology shock has a long-run effect on productivity growth, i.e. the long-run responses of productivity growth to all but one shock are restricted to zero.

The estimation stage then proceeds by finding those parameter values such that the model impulse responses match the empirical impulse responses (to each of the identified structural shocks, separately), for output, inflation, real wages and the nominal interest rate. The weighting matrix used for this minimization is a diagonal matrix with the sample variances of the impulse responses along the main diagonal. This is repeated one thousand times using bootstrapped samples of the SVAR and impusle responses, so that the distribution of the parameter estimates may be determined without relying on the asymptotic distribution, which may not be appropriate. The evaluation metrics of the models are the distributions of the root mean square error (RMSE) of the model impulse responses. It is found that the RMSE in relation to a monetary policy shock is almost identical for both model variants, but in relation to a technology shock the sticky price model does better in the sense of a lower mean RMSE and a narrower RMSE distribution. It is noteable that the computed confidence intervals for the estimates of

[9]The use of which is described, for example, by Lutkepohl (2006), chapter 9.

the updating probabilities in the sticky information model are quite wide, being between 0.3 and 0.5 in width. Carrillo (2012) also estimates a 'dual stickiness' model, finding that this model performs better than the pure sticky information model and about as well as the sticky price model, in relation to a technology shock. The estimates for this model vary quite a lot depending upon which shock the responses are matched upon and the confidence intervals in both cases are quite wide. This is suggested as a sign of potential partial identification, which is often a problem in the more general SVAR literature (see, for example, Canova (2007), chapter 4) . Following these results, several further variations are estimated. The first variation is the extension of the model to include capital, by adding an 'investor' to the household, but this agent is assumed to have full information. Allowing for this both the sticky price and sticky information versions perform about as well as each other, however the improvement in performance may be indicative of the general effect of including capital in the model rather than revealing any particular deficiency in the sticky information model. A second variation is the use of a data sample beginning in 1979. Now the estimates are mostly similar to the whole sample estimates, except that for the sticky price model the monetary policy coefficient on inflation is higher than for the whole sample, whereas for the sticky information model the monetary policy coefficient on inflation is unchanged, but the coefficient on the output gap is higher, relative to the whole sample results. Overall for this sub-sample the RMSE distributions concerning a monetary policy shock are similar for both models, while the sticky price model continues to do better concerning a technology shock.

In total this section has described the results from six different attempts to estimate various models featuring sticky information. To the extent that any general conclusions are possible given the variety of approaches used and some of the potential issues raised concerning the reliability of the results, the evidence seems to suggest that there is some support for the inclusion of sticky information as a general model feature, but that it does not necessarily perform any better as a single explanation of persistence/rigidity than sticky prices. All of these studies featured updating probabilities which were exogenously given; the results in section 4.4 extend this literature by allowing for endogenously-determined updating probabilities.

4.3 Estimation Methods

The first estimation method used here is indirect estimation (introduced by Smith (1993) as 'Extended Method of Simulated Moments'). This method proceeds as follows. Firstly, an unrestricted VAR(p) is estimated on the sample data, giving estimated coefficient matrices. For example, for a VAR(1):

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + u_t \quad \rightarrow \\ Y_t &= \hat{A}_1 Y_{t-1} + \hat{u}_t \end{aligned} \tag{4.1}$$

where ˆ denotes an estimated quantity, Y a vector of observed data, u a vector of residuals of equal dimension and A a square matrix of coefficients. Then, provided that the theoretical model can be simulated for given fundamental parameter values (which is the case here) a simulated data series can be produced and a VAR(1) (for example, but the method is the same for a VAR of order p) estimated on this simulated data. For the model solution in equation 3.25:

$$\begin{aligned} Z_t &= \sum_s \sum_{n=0}^{\infty} \hat{z}_n(s) e^s_{t-n} \\ &= \hat{A}(\zeta) Z_{t-1} + \hat{u}_t \end{aligned} \tag{4.2}$$

where ζ is the set of fundamental parameters used to solve the model and $\hat{A}(\zeta)$ is the matrix of coefficients estimated on the simulated data. The estimated fundamental parameters are then found as:

$$\zeta^* = argmin_\zeta \, vec(\hat{A} - \hat{A}(\zeta))' \, W \, vec(\hat{A} - \hat{A}(\zeta)) \tag{4.3}$$

The criterion for estimation is therefore that the model is able to produce simulated data which results in VAR coefficients which match those of the data. As the $\hat{A}(\zeta)$ are estimated using simulated data, they depend on the draws of exogenous shocks used in the simulation. In order to reduce the random effects of this on the estimates the simulation is repeated several times and the VAR coefficients are averaged to give the simulated model coefficients as used in the minimization, i.e.:

$$\hat{A}(\zeta) = \frac{1}{N} \sum_{n=1}^{N} \hat{A}_n(\zeta) \tag{4.4}$$

where n denotes the n-th set of exogenous shocks and a total of N different sets of such shocks are used. Under certain conditions, as described in Smith (1993), this method will produce consistent estimates of the fundamental parameters.

An advantage of this method is that it avoids the problems of identifying structural shocks which occur in structural vector auto-regression modelling and the inability to estimate the parameters of the structural shocks (due to the normalization of the structural covariance matrix which is necessary to identify the shocks, as will be discussed below). Thus indirect estimation is able to produce estimates related to the data for any model which can be simulated without it having to be cast into a form amenable to direct estimation. This is particularly useful here given that the model is not of the usual state-space form. Using indirect estimation is also relatively ‘cheap’ computationally as the simulations can be conducted in a few seconds, at most.

A problem with applying these methods to the model here is that once a set of estimated parameters has been obtained there is an additional iterative step of computing the updating probability equilibrium given these parameter values, then, because these probabilities are unlikely to be the same as those upon which the estimation was based, repeating the estimation subject to these new probabilities and so on until the full set of parameters is mutually consistent and represents an equilibrium of the endogenous sticky information model. This greatly increases the computation time required, to the order of several hours or even days, so it becomes impractical to then bootstrap the distribution of the estimates, as is commonly done with these methods. Therefore the precision of the estimates cannot be evaluated, although consistency should ensure that the results are still meaningful.

An additional problem related to this is that normally the EMSM method would proceed by initialising the minimization of the moment distances with an identity matrix as the weighting matrix. Then once the first set of estimates was obtained the implied sample variances would be used to form an estimate of the optimal weighting matrix, which would then be used to re-estimate the parameters. This process would be iterated until the parameter estimates and the estimate of the optimal weighting matrix converged. However, due to the iterative step introduced here by the inclusion of endogenous updating probabilities and the resulting increase in computation time, it would be impractical to repeat this for iterations of the estimation of the weighting matrix also. Therefore the results here are produced simply for the identity weighting matrix. This has the effect of attaching equal weight to each of the moments used for estimation[10], which is desirable, however it means that the estimates will not be efficient, only consistent. For the reasons given in Smith (1993) it would also not be appropriate to use the numerical asymptotic standard errors of the estimates, due to the use of simulation. This possible inefficiency must be kept in mind when considering the results in section 4.4.

The second empirical method used here is that of SVAR with identification of structural shocks by sign restrictions. The method here follows that proposed by Canova and De Nicolo (2002)[11]. Specifically, suppose that a first-order VAR was estimated on the data, giving:

$$Y_t = \hat{A}_1 Y_{t-1} + \hat{u}_t \tag{4.5}$$

The estimated residuals here, $\hat{u}_t$, are not the structural shocks specified in the theoretical model, so the impulse responses derived from this estimation would not be economically meaningful. However, if it is possible to decompose the estimation residuals into structural shocks then the impulse responses will be useful. Using the MA representation of this VAR

$$Y_t = B(L)u_t = \sum_{i=0}^{\infty} A_1^i u_{t-i} \tag{4.6}$$

[10]For example, Mankiw and Reis (2006a) use an identity weighting matrix for this reason.
[11]The same notation is followed where possible.

Given the residual covariance matrix Σ_u, the eigenvalue-eigenvector decomposition

$$\Sigma_u = PDP' \tag{4.7}$$

will produce an orthogonal decomposition of the covariance matrix. Defining $V = PD^{0.5}$, $\Sigma_u = VV'$. Therefore the MA representation of a structural model specified as

$$Y_t = C(L)V^{-1}e_t \tag{4.8}$$

will have the same covariance matrix as the unrestricted VAR, assuming that the structural shocks have an identity covariance matrix,$\Sigma_e = I$ and $C(L) = B(L)V$. The problem of identification is therefore to find a decomposition of the residual covariance matrix such that the resulting transformed shocks can be interpreted as the theoretical structural shocks, with unit variance. The means of doing so proposed by Canova and De Nicolo (2002) is to take advantage of the fact that the eigenvalue-eigenvector decomposition can then be transformed further using orthogonal matrices. In particular, they use Givens rotation matrices, defined as an identity matrix conditional on indexes m,n and rotation angle θ, where element $m, m = cos(\theta)$, element $m, n = -sin(\theta)$, element $n, n = cos(\theta)$ and element $n, m = sin(\theta)$. Such matrices are orthogonal, so the decomposition continues to be admissible

$$\begin{aligned} \Sigma_u = VV' = PD^{0.5} \\ = VV' = PD^{0.5}G_{m,n}(\theta)G_{m,n}(\theta)'D^{0.5}P' \end{aligned} \tag{4.9}$$

and the structural representation now takes the form

$$\begin{aligned} C(L) = B(L)PD^{0.5}G_{m,n}(\theta) \\ e_t = (PD^{0.5}G_{m,n}(\theta))^{-1}u_t \end{aligned} \tag{4.10}$$

Using this method it is possible to search over a range of values of the rotation angle and variables to rotate until a decomposition is found which gives impulse responses (i.e. the coefficients of the MA representation) that satisfy the restrictions appropriate to the signs of the responses to the structural shocks. If several decompositions satisfy the restrictions then the results which follow from them may be compared. If no decomposition satisfies all of the sign restrictions then alternative restrictions may need to be considered, or the decomposition which satisfies the greatest number of restrictions may be used, but this limited identification may bias the results. The particular sign restrictions used for identification here are given in section 4.4.2 alongside the results. The purpose of using this method is not to estimate the parameter values but to provide a check on the results from the indirect estimation. As discussed above, the addition of endogenous sticky information to the model extends the computation time such that some of the usual assessments of the estimates are impractical, so as an alternative the identified empirical impulse responses will be compared to the impulse responses of the model given the parameter estimates, to see how well the estimated model can fit the empirical responses.

The full sample used for estimation covers the period 1955 Q1 to 2006 Q1 and estimates are also produced for the two sub-samples 1955 Q1:1984 Q3 (the 'early' sub-sample) and 1984 Q3:2006 Q1 (the 'late' sub-sample). The sub-sample estimations are conducted because it is widely considered that there may have been a structural break in the USA economy (and in monetary policy in particular) around this point - whether this 'Great Moderation' was a consequence of changes in the properties of the exogenous disturbances affecting the economy or due to the effects of the change in monetary policy initiated by the change of Federal Reserve chairman in 1979 remains disputed, but it is sufficient here to accept this as a break date then consider any changes in parameter estimates which arise. Branch et al. (2009) use 1984 as the appropriate break data, so the results here ought to be comparable. Also, the full sample covers the same date range as used by Mankiw and Reis (2006b). The particular data series used were, for the non-farm business sector, real output per person, implicit price deflator, average hours per person, nominal hourly wage, real hourly wage and the effective federal funds rate. The data was obtained from the FRED2 database maintained by the Federal Reserve Bank of St Louis. All series are quarterly and seasonally adjusted, except the effective federal funds rate, which is available as a monthly average and was converted to a quarterly equivalent. These series were then used to produce the variables used in the estimations, with detrending carried out using the Hodrick-Prescott filter.

4.4 Results

4.4.1 Indirect Estimation Results

The first set of parameter estimates, for the whole data sample, is presented in table 4.1 below. The estimated updating probabilities are slightly lower than those used in the benchmark numerical calculations in chapter 3, especially for firms. The estimated monetary policy coefficients suggest a more responsive policy with respect to both inflation and the output gap than in the benchmark case. It is also noticeable that the standard deviation of the policy shock is much greater than that previously considered. The estimates of the goods and labour substitution elasticities are similar to those found by Mankiw and Reis (2006b).

Parameter	Value	Parameter	Value	Parameter	Value
λ	0.538	$\rho_{\Delta a}$	0.37	ρ_g	0.13
ω	0.171	$\sigma_{\Delta a}$	0.022	σ_g	0.011
γ	4.03	ρ_γ	0.89	ρ_ε	0.41
ρ_ν	0.79	σ_γ	0.008	σ_ε	1.1
σ_ν	1.83	ϕ_p	1.98	ϕ_y	0.67
ν	34.0				

Table 4.1: Indirect Estimation Results - Full Sample

Comparing these full sample results to the results for the 'early' sub-period shown in table 4.2, it can be seen that the updating probability of firms is very similar, while the updating probability of households is approximately one-third lower. The estimates of the monetary policy coefficients suggest that policy was much less responsive to inflation and the output gap during this period (which is in keeping with other findings) and that shocks to monetary policy were much less variable. Overall the change in the variability of shocks appears to have outweighed the effect of less responsive policy and induced households to update less frequently on average. However, firms updated roughy

Parameter	Value	Parameter	Value	Parameter	Value
λ	0.525	$\rho_{\Delta a}$	0.36	ρ_g	0.89
ω	0.11	$\sigma_{\Delta a}$	0.24	σ_g	0.001
γ	3.79	ρ_γ	0.52	ρ_ε	0.93
ρ_ν	0.91	σ_γ	0.001	σ_ε	0.02
σ_ν	1.93	ϕ_p	1.09	ϕ_y	0.24
ν	33.99				

Table 4.2: Indirect Estimation Results - Early Sub-Sample

as often on average, which suggests that there may also have been an effect from the volatilities of the substitution elasticities. σ_γ decreased substantially, perhaps allowing households to reduce ω[12], while σ_ν was approximately constant, perhaps causing firms to maintain the level of λ. As several of the parameter estimates have changed it is not immediately obvious which of these effects might be stronger, but overall the estimates are comprehensible in view of the model.

Table 4.3 shows the estimated parameters for the 'late' sub-period. The prior expectation about these results, informed by the narrative of the 'Great Moderation', might be that shocks would have become less variable and/or policy would have become more responsive, especially to inflation, both of which would allow agents to reduce the average frequency with which they updated their information. It can be seen from the table that monetary policy shocks were less variable, although not by very much, and monetary policy was more responsive to inflation and less responsive to the output gap. The estimated updating probabilities are $\lambda = 0.179$ and $\omega = 0.202$. It is understandable that households would increase their updating probability if they are mostly concerned with output and wages, as policy became less concerned with the output gap and the volatility of shocks to the wage mark-up (through γ) increased. The reduction in the updating probability of firms is also explicable as a consequence mainly of the increase in policy responsiveness to inflation, in accordance with the theoretical results in chapter

[12]Although this does not entirely match the numerical results in figure 3.4

Parameter	Value	Parameter	Value	Parameter	Value
λ	0.179	$\rho_{\Delta a}$	0.094	ρ_g	0.99
ω	0.202	$\sigma_{\Delta a}$	0.83	σ_g	0.45
γ	4.04	ρ_γ	0.80	ρ_ε	0.98
ρ_ν	0.60	σ_γ	0.013	σ_ε	0.018
σ_ν	1.76	ϕ_p	1.45	ϕ_y	0.15
ν	34.0				

Table 4.3: Indirect Estimation Results - Late Sub-Sample

3, although the value of the probability is lower than was typically found by the other studies (which did not allow for endogenous sticky information) discussed in section 4.2.

4.4.2 Alternative filtering/detrending

The results above were based on the use of the Hodrick-Prescott filter to detrend the data. It is arguable, however, that this is not the most appropriate way in which to do so. Firstly, the HP filter determines the trend at each date allowing for future values of the variables - this is not consistent with a model focussed on expectations and the dating of information, as the deviations from trend produced by the filter will not be comparable with the deviations from steady-state values as defined in the model. Secondly, univariate filtering takes no account of the relationship between the different variables (as noted by Fukac and Pagan (2010). This second criticism is also applicable to the alternative of differencing those variables which may be non-stationary - if each of the non-stationary variables possesses an independent unit root then differencing would be appropriate, but if there is cointegration between the variables and therefore one or more common stochastic trends, differencing would not be appropriate. Garratt et al. (2006) argue on these grounds that the appropriate detrending method in the presence of cointegration is the multivariate Beveridge - Nelson decomposition and they show how this can be derived from the cointegrated VAR/VECM representation, with the trend given by the conditional long-run forecast at each date. An advantage of the Garratt et al. (2006) approach is that the basis of the decomposition is the cointegrating relations between the variables, which potentially allows for an economic justification of the decomposition rather than a purely statistical ('black box') justification. Specifically, Garratt et al. (2006) show that for a VECM

$$\triangle x_t = g + \alpha(\beta^{'} x_{t-1} - \kappa) + \Phi(\triangle x_{t-1} - g) + e_t \tag{4.11}$$

the transitory components are given by

$$x_t - \hat{x}_t = -\alpha_{\infty}(\beta^{'} x_t - \kappa) - \Phi_{\infty}(\triangle x_t - g) \tag{4.12}$$

where $\hat{x}_t$ is the permanent component and a subscript of ∞ denotes the matrix of coefficients related to the long-run forecast.

The EViews code necessary to implement this method is available from the authors[13] and this was used to apply the decomposition to the USA data sample.

Two of the variables in the sample used here, hours and the nominal interest rate, might be expected theoretically to be stationary but this seems to be unclear in the actual data. Therefore a VAR in all five variables was estimated, with the lag order 'selected' by the AIC being 2, then a Johansen trace test was conducted. The result of this test was that there were two cointegrating relationships in the data - if hours and the interest rate were stationary then this might imply one common stochastic trend amongst the three remaining non-stationary variables, which would be consistent with the model of chapter 3, in which the common trend originated in the technology process.

For the decomposition used here the restrictions applied to the cointegrating equations were those required for exact identification rather than any over-identifying restrictions motivated by economic theory, although it may be worth attempting to test such restrictions in future work. Given the detrended data series obtained via this decomposition, the indirect estimation of section 4.4.1 was repeated, with the results for the earlier and later sub-samples reported in tables 4.4 and 4.5 below.

[13] At www.econ.bbk.ac.uk/faculty/wright

Parameter	Value	Parameter	Value	Parameter	Value
λ	0.620	$\rho_{\Delta a}$	0.33	ρ_g	0.89
ω	0.243	$\sigma_{\Delta a}$	0.12	σ_g	0.02
γ	3.84	ρ_γ	0.69	ρ_ε	0.87
ρ_ν	0.648	σ_γ	0.34	σ_ε	0.014
σ_ν	2.25	ϕ_p	1.23	ϕ_y	0.39
ν	32.31				

Table 4.4: Alternative IE Results - Early Sub-Sample

Parameter	Value	Parameter	Value	Parameter	Value
λ	0.685	$\rho_{\Delta a}$	0.26	ρ_g	0.98
ω	0.435	$\sigma_{\Delta a}$	0.036	σ_g	0.008
γ	4.40	ρ_γ	0.76	ρ_ε	0.90
ρ_ν	0.92	σ_γ	0.02	σ_ε	0.41
σ_ν	1.69	ϕ_p	1.20	ϕ_y	0.26
ν	35.56				

Table 4.5: Alternative IE Results - Late Sub-Sample

It can be seen from these tables that the results are not entirely independent of the detrending/filtering method used. The results now suggest that firms increased their updating probability by a small amount and households increased their updating probability substantially. As previously, the behaviour of households could perhaps be explained primarily by the decrease in the responsiveness of monetary policy to the output gap. In contrast to the results for the HP data, both σ_ν and σ_γ have decreased, while σ_ε has increased considerably (which may be why there is no decrease in firm updating in this case). These results are plausible in that the estimated changes in the information acquisition behaviour are compatible with and explicable in relation to the other parameter changes. However, the results are not entirely consistent with those of section 4.4.1, which may not be surprising given the different detrending method used, but this suggests that neither set of results can be taken as definitive. Nonetheless, the results could be considered to collectively offer some support for the incorporation of sticky information, in that the changes between sub-periods are mutually compatible with the estimated changes in the exogenous variables.

4.4.3 SVAR impulse responses

Table 4.6 shows the sign restrictions used to identify the structural shocks in the SVAR, which are based on the signs of the theoretical model impulse responses, i.e. the computed impact responses under the benchmark parameterization of chapter 3. These restrictions could of course be compatible with other models also, especially given that none are too counter-intuitive, so using these restrictions for identification should not pre-bias the identified impulse responses towards matching those of the model under the estimated parameters. It can be seen from table 4.6 that the restrictions are that: an expansionary monetary shock should have a positive effect on all variables except for the nominal interest rate; a positive technology growth shock should increase output growth and real wage growth but reduce inflation, hours and the nominal interest rate; a positive aggregate demand/fiscal policy shock should increase all variables except for real wage growth; a positive shock to the substitution elasticity of labour in production (equivalent to a negative shock to the nominal wage markup) should increase output and hours, while reducing inflation, the nominal interest rate and real wage growth; and finally that a positive shock to the substitution elasticity of goods in consumption (equivalent to a negative shock to the goods price markup) should reduce inflation and the nominal interest rate, while increasing output growth, real wage growth and hours. Again, these response signs all seem quite plausible and not necessarily specific to the model used here, so the impulse responses in the data to the shocks identified by these restrictions are not being forced to comply with this particular model. Note that the restrictions in table 4.6 are for the impact impulse responses (the effect in period t); the actual identification procedure also used the signs of the effects in period 1 - these are not shown here because all of the signs are the same as for period t, except for the response of real wage growth to a monetary policy shock and the response of output growth to a demand shock.

	ε	Δa	g	γ	ν
π	+	-	+	-	-
Δy	+	+	+	+	+
l	+	-	+	+	+
i	-	-	+	-	-
Δrw	+	+	-	-	+

Table 4.6: SVAR sign restrictions

Table 4.6 also shows that the structural shocks are in principle identifiable, in the sense that none of the shocks share the exact same sign restrictions across all of the variables, so for a set of identified shocks satisfying these restrictions the assignment to the theoretical structural shocks is unambiguous.

Figure 4.1 shows the impulse response functions for the model of chapter 3 given the estimated parameters of section 4.4.2 (given that the detrending method used in that section had the greatest theoretical justification) and the impulse response functions of the (full-sample) data to the sign-restriction-identified structural shocks. The first column shows the responses to a monetary policy shock, the middle column shows the responses to a technology/productivity shock and the third column shows the responses to a government spending/demand shock. The responses are shown for two years following the occurrence of a shock. The first row of the figure shows the responses of inflation to the different shocks, the second row shows the responses of output growth to the different shocks and the third row shows the responses of real wage growth to the different shocks. The legend shown in the left-hand plot of the middle row is applicable to all of the plots.

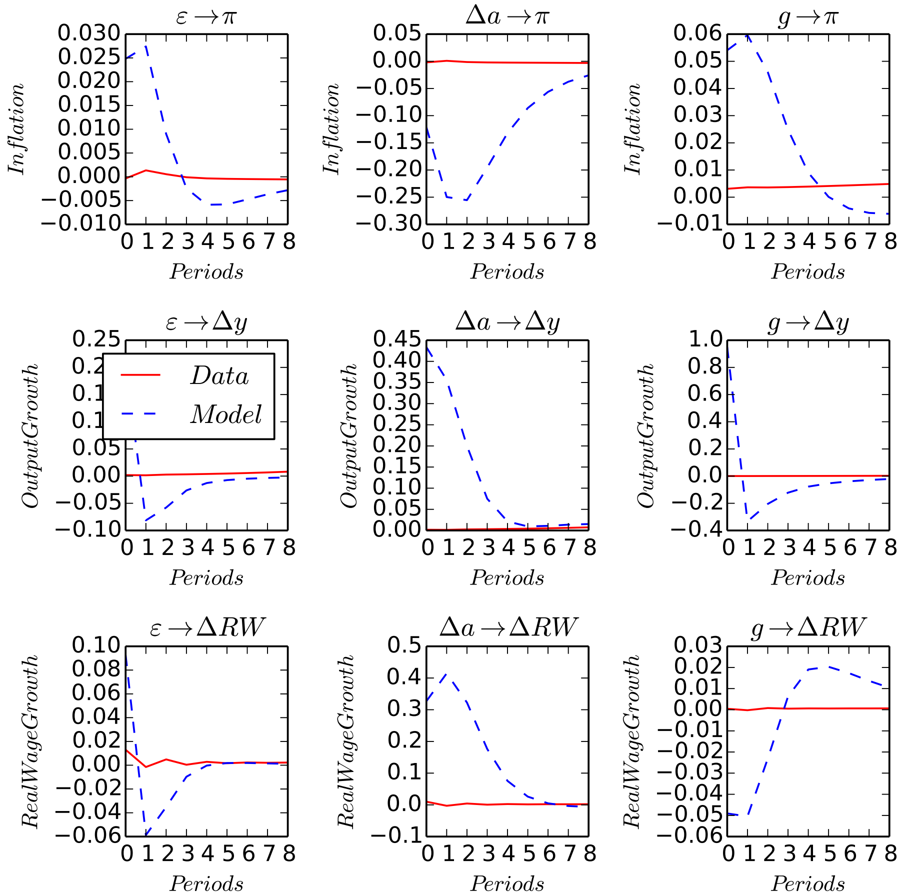

Figure 4.1: Model and Structural Impulse Response Functions

Considering firstly the top row of figure 4.1, it can be seen that the model given the estimated parameters does not match very closely the empirical impulse responses of inflation to any of the three shocks presented. In each case the model shows a much larger and much more persistent response than the data. From the second row of the figure it is apparent that the model does not match the empirical responses well for output growth. The model suggests that there is a smooth decrease over time in the effect of a monetary policy shock on output growth, but the empirical response does not agree with this. The third row of figure 4.1 shows little similarity between the model and empirical responses of real wage growth to any of the shocks, apart perhaps from the sign of the impact response. As noted previously, a difficulty in assessing these results is the absence of confidence bands, which prevents the comparison of the distribution of responses and therefore may make the disparity between the model and empirical impulse responses seem greater. Of course, it is possible that the two sets of responses may still not match very closely even were it possible to consider the distribution. Indeed, the empirical responses all seem to be extremely small, such that it might be difficult to accept these as indicative of the true empirical responses.

In order to see whether the disparities evident in figure 4.1 may be due to a lack of identification, figure 4.2 plots the model impulse responses against the orthogonalized impulse responses from an unrestricted VAR estimated on the data. While the plots in this figure do show the responses of the named variables, the impulse name can only be loosely interpreted, e.g. the ε impulse in this figure is actually just the orthogonalized shock to the interest rate equation in the VAR, but is not properly structurally identified (given that the model does not appear to support the structural form implied by a Choleski factorization).

Figure 4.2 suggests that several of the model impulse responses do match the signs of the empirical impulse responses, but fail to match the scale or dynamics of the empirical responses. To the extent that one of the original reasons for the interest in the sticky information model was its ability to produce realistic impulse responses, this might be thought to be a form of refutation. However, the impulse responses produced by the model do tend to match the 'stylized' responses frequently found in the literature, so it could perhaps be argued that it is unlikely that either of the approaches used here has successfully identified the empirical impulse responses.

The comparisons in figures 4.1 and 4.2 do not offer a great deal of support for the endogenous sticky information model. There are several factors which may mitigate this. Firstly, it is possible that the structural shocks are not fully identified, despite the satisfaction of the majority of the sign restrictions and the consideration of those shocks which did satisfy the restrictions. In the model the shock processes are specified exactly, while the identification of the empirical shocks by sign restrictions does not necessarily

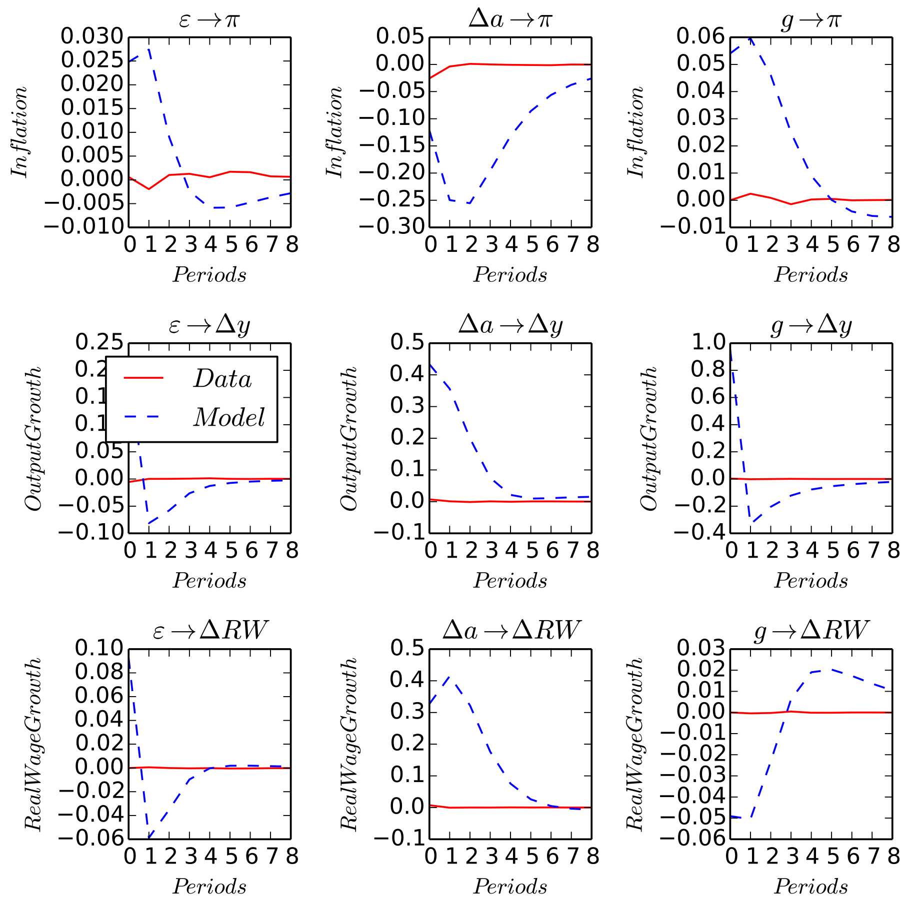

Figure 4.2: Model and Unrestricted Impulse Response Functions

impose a similar process on the shocks, so the remaining parameter equivalents in the data may be very different from those used for the model. As noted by Fry and Pagan (2011), when the unrestricted data VAR has a selected lag order of 1 the use of sign restrictions related to impulses beyond period t + 1 does not actually add any more to the ability of the procedure to identify the structural shocks, as the restrictions are merely the sign of increasing powers of the VAR coefficients. Therefore it is not possible to improve the identification here using more restrictions, unless the VAR lag selection criteria are ignored, which would not improve the representation of the data or be likely to improve the results overall. Secondly, it is possible that the parameter estimates used to generate the model responses are not very precisely estimated, which would reduce the ability of the model to fit these characteristics of the data, although this might be a less likely explanation than the first explanation above as the estimates have at least been consistently estimated.

4.5 Conclusion

The results in this chapter show that the endogenous sticky information model is able to explain several features of the USA data, although this conclusion must be somewhat tentative. The estimates of the updating probabilities in different periods show a quite large change, with the more recent period having less frequent updating by firms and slightly more frequent updating households. These changes are explicable by the changes in the other parameter estimates between the two sub-periods. In particular, it has been found that monetary policy has become stricter concerning inflation and less concerned about the output gap. This has been observed elsewhere previously, however the result that this influences the sticky information probabilities, in the manner which the theoretical model would predict, is new. Similarly the result that the changes in the price and wage mark-up volatilities lead to theory-consistent changes in the updating probabilities is also new. The sensitivity of the results to alternative detrending was also considered, with it being found that the individual parameter changes were quite sensitive to the detrending method used, although in both cases there did seem to be some support for sticky information in that the parameter changes were consistent with the theoretical/simulated results of chapter 3.

However, the incorporation of endogenous sticky information into a DSGE model cannot be claimed to improve the general fit to the data. As shown by the comparison of the estimated model impulse response functions to the empirical structural impulse responses, the model struggles to match the dynamics exhibited within the data. As noted, it is not possible to determine conclusively whether this is a consequence of partial identification of the structural shocks in the data, or due to some other misspecification of the model which is preventing the model from replicating the features of the data.

Chapter 5

Conclusion

This thesis has provided several new results concerning the endogenous sticky information model. Beginning with the simple model of monopolistically competitive firms with exogenous aggregate demand considered in chapter 2, it was found that for a plausible degree of strategic complementarity in price-setting the model could have multiple equilibria and, in particular, there were multiple equilibria of the updating 'game' between firms. It was also found that the choice of updating probability depended quite strongly on the costs of information and the variability and persistence of the exogenous shocks, with the relationships being those which would be expected intuitively. When the possibility of changes in the volatility of the exogenous monetary policy shock was considered, with the interpretation that such changes could be considered to be changes in the 'success' with which monetary policy followed a rule, it was found that there was a trade-off between reducing/increasing the variance of output and increasing/reducing the variance of inflation. This trade-off does not occur in the model with exogenous sticky information, partly because there is no possibility of multiple equilibria in that model. This is a significant finding, in that it demonstrates that there may be additional implications for policy changes when agents respond to such changes by altering the probability with which they receive new information. However, the applicability of this finding may be limited by the relative simplicity of the model, which does not contain many features which would be considered desirable in a macroeconomic model that was being used to consider monetary policy.

Therefore the third chapter considered a DSGE model more in accordance with current theory. The particular model followed that of Mankiw and Reis (2006b), but with the extension of endogenous sticky information. In this case there were two different types of agent, firms and households, each with a separate updating probability. In the model price-setting firms were monopolistically competitive and hired labour from wage-setting trade union/households. Monetary policy was set according to a fairly standard Taylor rule specification. Additional exogenous shocks to the model came from aggregate demand/fiscal policy, the wage mark-up and the price mark-up. It was found that the optimal updating probabilities of different agent types displayed an asymmetric strategic inter-relationship, with firms responding quite strongly to the choice of households but households responding quite weakly to the choice of firms. Perhaps as a consequence of this, this model does not appear to possess multiple equilibria, at least under the range of parameterizations investigated. It was also found that the optimal updating probabilities varied with the costs of information, the standard deviations of the exogenous shocks and the coefficients of the monetary policy rule.

Numerical experiments with this model showed that when there was an increase in the responsiveness of monetary policy to inflation (with an assumed target of zero) the effect varied between exogenous and sticky information models. For exogenous sticky information it was found that an increase in ϕ_p reduced the standard deviations of both output and inflation. However, when the same change was considered under endogenous sticky information it was found that the policy change was amplified with respect to inflation, i.e. the change in the inflation standard deviation was greater than in the exogenous sticky information case, but with respect to output the effect of the policy change was dampened, with a smaller decrease in output variability. For changes in the volatility of monetary policy shocks, σ_ε, it was found that again the presence of endogenous sticky information had a significant effect. As the volatility increases the updating probabilities of both types of agents increase, and in the case of firms the optimal choice was actually to be constantly fully-informed. In the case of exogenous sticky information, when σ_ε increases the variances of output and inflation increase, and when endogenous sticky information is allowed for both of these variances increase by a larger amount, as the increase in updating probabilities means that prices, wages and consumption respond much more quickly to the effects of shocks. Hence any change in monetary policy which corresponded to a reduction of the accuracy of the Taylor rule as a characterization of policy would be amplified with respect to the variability of output and inflation. Concerning the dynamics of the economy, the results of the numerical simulation of the impulse response functions of output and inflation in chapter 3 revealed two aspects of the effects of monetary policy changes. Firstly, an increase in the responsiveness of monetary policy to inflation results in agents updating less frequently on average, which results in both inflation and output responding less strongly to shocks, with smaller maximal responses and generally smaller responses for several quarters after the impact date. A second, and perhaps more unusual, result was that a change in endogenous updating probabilities, induced by a change in monetary policy coefficient, could change the length of time before the maximal response occurred (that is, the peak of the impulse response function occurred at a different date). While this effect was not very strong in terms of the shift of time, combined with the change in the size of responses this was seen to lead to quite different dynamics compared to the benchmark case.

Chapter 3 also considered the dynamics of output and inflation in response to a monetary policy shock when the standard deviation of the shock changed. In the exogenous sticky information case such a change would merely rescale the impulse response function. However, it was found that in the endogenous sticky information case both the size of the response (beyond the rescaling) and the shape of the response changed. For a reduction in σ_ε both agents chose to update less frequently and as a result the responses of output and inflation to shocks were smaller and, in the case of output, had a later peak than the benchmark response. For an increase in σ_ε both agents chose to update more frequently and consequently the immediate responses of output and inflation to a shock were much greater and the peaks of the response functions occurred sooner.

Collectively, the numerical results obtained in chapter 3 suggest that the addition of endogenous sticky information to a model is highly significant, with very different properties of output and inflation. To the extent that this specification of the informational limitations faced by agents when forming their expectations is accurate, these results suggest that the incorporation of such a feature into models is essential, as its omission would not be innocuous (this may be especially true for a model in which policy is set optimally).

The estimation results of chapter 4 showed changes in the updating behaviour of private agents within the USA economy over the period 1955 - 2006 and suggested some of the possible reasons why this may have occurred. Using the extended method of simulated moments/indirect estimation to match the coefficients of a VAR estimated on simulated data with the coefficients of a VAR estimated on the actual data, it was found that relative to the period prior to 1984, firms have since chosen much lower updating probabilities while households have chosen much higher updating probabilities. This finding is quite surprising, but the results also indicate why this may have happened. The first potential explanation is that monetary policy became more responsive to inflation and less responsive to the output gap, which may have reduced the exposure of firms to variability of prices but increased the exposure of households to variability of consumption. A second possible explanation is that the estimation results show a slight decrease in σ_ν and a large increase in σ_γ. These would lead to less volatile price mark-ups but more volatile wage mark-ups, allowing firms to update less frequently without experiencing large expectational errors, but requiring households to update more frequently in order to reduce the increased expectational errors they might face. These two explanations are complementary and accord with the theoretical results found in chapter 3.

Chapter 4 also considered the ability of the estimated model to match some of the dynamic responses of variables to exogenous shocks. This was done using the method of SVARs with identification of the structural shocks by the imposition of sign restrictions on the impulses. These results did not suggest that the estimated model would perform very well empirically, as few of the empirical impulse responses were closely matched by the estimated model responses. Some of the possible reasons for this which were considered were that the identification method may have failed to fully identify the empirical counterparts of the structural shocks and that the comparison may have been somewhat biased towards 'rejection' because it was not possible to produce confidence bands for the impulse responses, which may have shown that there was at least some overlap with the distribution of the responses, even thought the 'point' estimates did not match.

Possible future theoretical work in the area of this research could look at heterogeneity of expectations and of information updating timing. Rather than the constant per-period probability of receiving new information used here, the model might be more realistic if individual agents could choose particular updating dates over time, with updating as a whole perhaps being modelled in a manner similar to Taylor staggered contracts. A time-varying choice of when to update would allow the various numerical experiments conducted with the model to reveal genuine dynamic behaviour, rather than comparative static results, which have the disadvantage of not showing the transition of the economy over time. The microeconomics of such a model has been analysed to some extent already, for example by the model of discrete choice updating of Drager (2011), however it has not been incorporated into a DSGE model in a fully satisfactory way. A possible empirical extension of this work would be to estimate the model of chapter 3 using data for the United Kingdom and other G-7 countries, as this would reveal whether there were any significant differences in information updating behaviour between different countries. Equally, it would be desirable to apply other estimation methods to this model, given the possible problems with the results as discussed in chapter 4. If it was possible to find a computationally feasible implementation of another method such as maximum likelihood then this could provide good supporting evidence for the wider inclusion of this type of information friction within DSGE models.

Appendix

A.1 Appendix to chapter 2

A.1.1 Individual Optimization

The derivations here closely follow those of Ball et al. (2005). Given the problem of maximizing 2.1 subject to the various constraints the first stage of the solution is for the agent to determine the consumption of each variety in order to minimize the cost of one unit of the aggregate consumption good:

$$min_{C_{it}^j} \int_0^1 P_{jt} C_{it}^j \, dj \; s.t. \; \left[\int_0^1 \left(C_{it}^j\right)^{\frac{\gamma-1}{\gamma}} dj \right]^{\frac{\gamma}{\gamma-1}} = C_{it} \tag{5.1}$$

which results in the first-order condition:

$$P_{jt} = \lambda_t (C_{it}^j)^{\frac{-1}{\gamma}} \left[\int_0^1 \left(C_{it}^j\right)^{\frac{\gamma-1}{\gamma}} dj \right]^{\frac{1}{\gamma-1}} \tag{5.2}$$

Re-arranging this and substituting out the Lagrange multiplier gives the individual demand for each variety as:

$$C_{it}^j = \left(\frac{P_{jt}}{P_t} \right)^{-\gamma} C_{it} \tag{5.3}$$

where the aggregate price index is

$$P_t = \left[\int_0^1 P_{it}^{1-\gamma} di \right]^{\frac{1}{1-\gamma}} \tag{5.4}$$

Given that all agents are identical, imposing the equilibrium condition that $C_t^i = Y_{it}$ and taking logs (denoted by lowercase variable names) gives the (log) aggregate demand function for good i as:

$$y_{it} = y_t - \gamma \left(p_{it} - p_t \right) \tag{5.5}$$

The second stage of this optimization is the agent setting the optimal price given the demand function for their output. This has first-order condition

$$\frac{P_{it}^*}{P_t} = \frac{\gamma}{\gamma - 1} \frac{C_t^{\sigma} L_{it}^{\psi}}{A_t} \tag{5.6}$$

Using the production function, the demand function (in levels) and the equilibrium between consumption and output, this can be expressed as:

$$\frac{P_{it}^*}{P_t} = \frac{\gamma}{\gamma - 1} \frac{Y_t^{\sigma} \left(\frac{Y_{it}}{A_t} \right)^{\psi}}{A_t} = \frac{\gamma}{\gamma - 1} \frac{Y_t^{\sigma} \left(\frac{\left(\frac{P_{jt}}{P_t} \right)^{-\gamma}}{A_t} \right)^{\psi}}{A_t} \tag{5.7}$$

Setting $A_t = \left(\frac{\gamma}{\gamma-1} \right)^{\frac{1}{1+\psi}}$ so that technology is not stochastic and the natural rate of output is normalized to zero, defining $\alpha = \frac{\sigma+\psi}{1+\psi\gamma}$ and taking logs of the above gives, after re-arrangement, the optimal price under flexible prices as:

$$p_{it}^* = p_t + \frac{log(\frac{\gamma}{\gamma-1})}{1 + \psi\gamma} + \frac{\sigma + \psi}{1 + \psi\gamma} y_t - \frac{(1 + \psi)}{1 + \psi\gamma} a_t = p_t + \alpha y_t \tag{5.8}$$

A.1.2 Solution for Expectations in Loss Function

Given the assumed stochastic processes for the disturbances substitution of equations 2.7, 2.8 and 2.9 into the optimal price, equation 2.4, results in:

$$\begin{aligned} p^*_{t+n} = (1-\alpha)\Bigg(\overline{\lambda} p^*_{t+n} + \overline{\lambda}\sum_{j=1}^{n-1}(1-\overline{\lambda})^j E_{t+n-j}\Big[p^*_{t+n}\Big] \\ +(1-\overline{\lambda})^n E_t\Big[p^*_{t+n}\Big]\Bigg) + \alpha\frac{\mu}{1-\delta} \\ +\alpha\sum_{h=0}^{\infty}\delta^h e_{t+n-h} + \sum_{k=0}^{\infty}\rho^k \varepsilon_{t+n-k} \end{aligned} \tag{5.9}$$

Then after gathering optimal price terms this is:

$$\begin{aligned} p^*_{t+n} = \frac{(1-\alpha)\,\overline{\lambda}}{1-(1-\alpha)\,\overline{\lambda}} \quad \textstyle\sum_{j=1}^{n-1}(1-\overline{\lambda})^j E_{t+n-j}\Big[p^*_{t+n}\Big] \\ +\frac{(1-\alpha)\,(1-\overline{\lambda})^n}{1-(1-\alpha)\,\overline{\lambda}} E_t\Big[p^*_{t+n}\Big] \qquad +\tfrac{\alpha\mu}{(1-(1-\alpha)\,\overline{\lambda})(1-\delta)} \\ +\frac{\alpha\sum_{h=0}^{\infty}\delta^h e_{t+n-h}}{1-(1-\alpha)\,\overline{\lambda}} + \qquad \tfrac{\sum_{k=0}^{\infty}\rho^k \varepsilon_{t+n-k}}{1-(1-\alpha)\,\overline{\lambda}} \end{aligned} \tag{5.10}$$

Taking the expectation of this conditional on time t information gives:

$$\begin{aligned} E_t\Big[p^*_{t+n}\Big] = \qquad \tfrac{(1-\alpha)\,\overline{\lambda}}{1-(1-\alpha)\,\overline{\lambda}} \qquad \sum_{j=1}^{n-1}(1-\overline{\lambda})^j E_t\Big[p^*_{t+n}\Big] \\ +\tfrac{(1-\alpha)\,(1-\overline{\lambda})^n}{1-(1-\alpha)\,\overline{\lambda}} E_t\Big[p^*_{t+n}\Big] \quad +\frac{\alpha\mu}{(1-(1-\alpha)\,\overline{\lambda})(1-\delta)} \\ +\tfrac{\alpha\sum_{h=n}^{\infty}\delta^h e_{t+n-h}}{1-(1-\alpha)\,\overline{\lambda}} + \qquad \frac{\sum_{k=n}^{\infty}\rho^k \varepsilon_{t+n-k}}{1-(1-\alpha)\,\overline{\lambda}} \end{aligned} \tag{5.11}$$

Gathering the expectations terms from equations 5.11 gives:

$$
\begin{aligned}
E_t\left[p^*_{t+n}\right] = \quad & \frac{\alpha\mu}{\left(1-(1-\alpha)\sum_{j=0}^{n-1}(1-\overline{\lambda})^j-(1-\alpha)\left(1-\overline{\lambda}\right)^n\right)(1-\delta)} \\
& \frac{\alpha\sum_{h=n}^{\infty}\delta^h e_{t+n-h}+\sum_{k=n}^{\infty}\rho^k\varepsilon_{t+n-k}}{1-(1-\alpha)\sum_{i=0}^{n-1}(1-\overline{\lambda})^i-(1-\alpha)\left(1-\overline{\lambda}\right)^n}
\end{aligned}
\tag{5.12}
$$

Then by similar re-arrangements the time t + 1 expectation can be expressed as:

$$
\begin{aligned}
E_{t+1}\left[p^*_{t+n}\right] = \quad & \frac{(1-\alpha)\left(1-\overline{\lambda}\right)^n}{1-(1-\alpha)\sum_{j=0}^{n-1}(1-\overline{\lambda})^j}E_t\left[p^*_{t+n}\right] \\
& +\frac{\alpha\mu}{\left(1-(1-\alpha)\sum_{i=0}^{n-1}(1-\overline{\lambda})^{ji}\right)(1-\delta)} \\
& +\frac{\alpha\sum_{h=n-1}^{\infty}\delta^h e_{t+n-h}+\sum_{k=n-1}^{\infty}\rho^k\varepsilon_{t+n-k}}{1-(1-\alpha)\sum_{l=0}^{n-1}(1-\overline{\lambda})^l}
\end{aligned}
\tag{5.13}
$$

In general the expectation of the time t + n optimal price conditional on time t + z information is[1]:

$$
\begin{aligned}
E_{t+z}\left[p^*_{t+n}\right] = \quad & \frac{(1-\alpha)\,\overline{\lambda}\sum_{j=1}^{z-1}(1-\overline{\lambda})^{n-j}E_{t+j}\left[p^*_{t+n}\right]}{1-(1-\alpha)\sum_{i=z}^{n}(1-\overline{\lambda})^{n-i}} \\
& +\frac{(1-\alpha)\left(1-\overline{\lambda}\right)^n}{1-(1-\alpha)\sum_{l=z}^{n}(1-\overline{\lambda})^{n-l}}E_t\left[p^*_{t+n}\right] \\
& +\frac{\alpha\mu}{\left(1-(1-\alpha)\sum_{m=z}^{n}(1-\overline{\lambda})^{n-m}\right)(1-\delta)} \\
& +\frac{\alpha\sum_{h=n-z}^{\infty}\delta^h e_{t+n-h}+\sum_{k=n-z}^{\infty}\rho^k\varepsilon_{t+n-k}}{1-(1-\alpha)\sum_{o=0}^{n}(1-\overline{\lambda})^{n-o}}
\end{aligned}
\tag{5.14}
$$

It is then possible to use this expression to replace for the expectations in the optimal price equation (equation 5.10). Using this and the above it is possible to then derive the following expression for the forecast error terms in equation 2.10:

$$
E_{t+z}\left[p^*_{t+n}\right]-p^*_{t+n} = \quad \sum_{b=0}^{n-z-1}-\frac{\alpha\delta^b e_{t+n-b}+\rho^b\varepsilon_{t+n-b}}{1-(1-\alpha)\sum_{j=0}^{b}(1-\overline{\lambda})^j}
\tag{5.15}
$$

Given that the two shocks are independent of each other, serially uncorrelated and homoscedastic, substitution of the above into the loss function results in equation 2.11 of the main text.

[1]The order of summation has been reversed to fit the change in dating expressions

A.1.3 Sign of General Cross-Partial Derivative of the Loss Function

The general cross-partial derivative with respect to strategies is:

$$
\begin{aligned}
\frac{\partial^2 Loss_i}{\partial\lambda_i\partial\overline{\lambda}} &= \frac{-2\beta\left(1-\alpha\right)\left(\alpha^2\sigma_e^2+\sigma_\varepsilon^2\right)}{(1-(1-\alpha)\overline{\lambda})^3} + \sum_{n=2}^{N}\beta^n\left[\sum_{z=1}^{n-1}(1-\lambda_i)^{n-z}\sum_{b=0}^{n-z-1}\left(\right.\right. \\
&\left.\left.\frac{\left(\alpha^2\delta^{2b}\sigma_e^2+\rho^{2b}\sigma_\varepsilon^2\right)\left(2\left(1-\alpha\right)\sum_{j=0}^{b}(1-\overline{\lambda})^j-2\left(1-\alpha\right)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)}{\left(1-(1-\alpha)\overline{\lambda}\,\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)^3}\right)\right] \\
&-\sum_{n=2}^{N}\beta^n\left[\lambda_i\sum_{z=1}^{n-1}\left(n-z\right)\left(1-\lambda_i\right)^{n-z}\sum_{b=0}^{n-z-1}\left(\right.\right. \\
&\left.\frac{\left(\alpha^2\delta^{2b}\sigma_e^2+\rho^{2b}\sigma_\varepsilon^2\right)\left(2\left(1-\alpha\right)\sum_{j=0}^{b}(1-\overline{\lambda})^j-2\left(1-\alpha\right)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)}{\left(1-(1-\alpha)\overline{\lambda}\,\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)^3}\right) \\
&+n\left(1-\lambda_i\right)^{n-1}\sum_{b=0}^{n-1}\left(\right. \\
&\left.\left.\frac{\left(\alpha^2\delta^{2b}\sigma_e^2+\rho^{2b}\sigma_\varepsilon^2\right)\left(2\left(1-\alpha\right)\sum_{j=0}^{b}(1-\overline{\lambda})^j-2\left(1-\alpha\right)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)}{\left(1-(1-\alpha)\overline{\lambda}\,\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)^3}\right)\right]
\end{aligned}
\tag{5.16}
$$

The first term in this expression is the same as for the one period problem, for which it has been shown that the sign depends upon the value of α. Then for the additional terms, beginning with the denominator of the bracketed fraction, $\left(1-(1-\alpha)\overline{\lambda}\,\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)^3$, which will have minimum value for $b=+\infty$: for $\overline{\lambda}=0 or 1$ this has a value of 1, otherwise given $\overline{\lambda}\in(0,1)\,,\,\sum_{j=0}^{\infty}(1-\overline{\lambda})^j=\frac{1}{1-(1-\overline{\lambda})}$

so $(1-\alpha)\overline{\lambda}\,\sum_{j=0}^{\infty}(1-\overline{\lambda})^j=\frac{(1-\alpha)\overline{\lambda}}{\overline{\lambda}}=(1-\alpha)$. Therefore the lowest possible value of this denominator term is α^3 which is positive for $\alpha>0$. Hence the sign of the bracketed fraction will depend upon the sign of the numerator. The first numerator term, $\left(\alpha^2\delta^{2b}\sigma_e^2+\rho^{2b}\sigma_\varepsilon^2\right)$, is clearly positive for parameter values within the permitted ranges. The second numerator term, $(2\,(1-\alpha)\sum_{j=0}^{b}(1-\overline{\lambda})^j - 2\,(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j)$, is equivalent to $2\,(1-\alpha)\,(b+1)\,(1-\overline{\lambda})^b$, the sign of which can be seen to depend on $(1-\alpha)$ - for $\alpha < 1$ this is positive (for finite b; for $b = +\infty$ this equals zero), whereas for $\alpha > 1$ this is negative. Therefore the overall sign of the bracketed fraction expression has the same relationship with the value of α.

Defining $\frac{\left(\alpha^2\delta^{2b}\sigma_e^2+\rho^{2b}\sigma_\varepsilon^2\right)\left(2\,(1-\alpha)\sum_{j=0}^{b}(1-\overline{\lambda})^j-2\,(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)}{\left(1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{b}(1-\overline{\lambda})^j\right)^3} = x$ and taking the sums for N $= +\infty$ (corresponding to the maximum absolute values of the sums) the above can be expressed (after simplification) as:

$$\frac{\partial^2 Loss_i}{\partial\lambda_i\partial\overline{\lambda}} = \frac{-2\beta\,(1-\alpha)\,(\alpha^2\sigma_e^2+\sigma_\varepsilon^2)}{(1-(1-\alpha)\overline{\lambda})^3} + \frac{(-\lambda_i^2\beta+\lambda_i^2\beta^2-2\lambda_i\beta^2+4\lambda_i\beta-2\lambda_i+\beta^2-3\beta+3)\beta^2 x}{(\beta-1)\,(\lambda_i\beta-\beta+1)^2} \tag{5.17}$$

It has already been established that the sign of 'x' depends on α. Recall that for $\alpha \in (0,1)$ x is positive, consequently

$$\frac{\partial^2 Loss_i}{\partial\lambda_i\partial\overline{\lambda}} = \begin{cases} <0 & if\; 0<\alpha<1 \\ =0 & if\; \alpha=1 \\ >0 & if\; \alpha>1 \end{cases} = \begin{cases} strategic\, complementarity \\ no\, strategic\, relationship \\ strategic\, substitutability \end{cases}$$

Figure 5.1 shows the value of the second term on the r.h.s in equation 5.17 for x = 1 (corresponding to some α positive but less than 1). It can be seen that this second term is negative, confirming the above result.

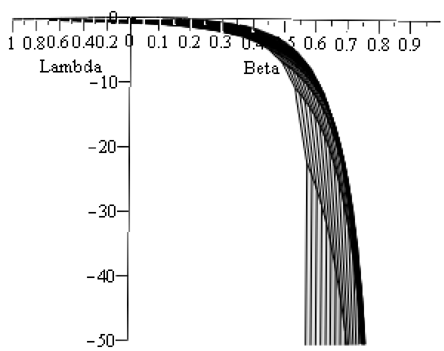

Figure 5.1: Value of second term of equation 5.17 for varying λ and β

A.1.4 Convergence of Infinite Horizon Loss Functions

For the loss function of equation 2.11 the approach here is to take all of the sums to infinity, then use the result that if $\sum_n b_n$ is convergent and $a_n < b_n \forall n$ then a_n is convergent also[2]. As there are several summations in equation 2.11 the following begins with the innermost sum (taken to infinity) and proceeds back to the complete sum:

First, take the denominator term in the main expected deviation expressions

$$1-(1-\alpha)\overline{\lambda}\sum_{j=0}^{\infty}(1-\overline{\lambda})^{j}=\begin{cases}\alpha & if\ 0<\overline{\lambda}\leq 1\\ 1 & if\ \overline{\lambda}=0\end{cases} \tag{5.18}$$

As $\overline{\lambda}=0$ is excluded by assumption the remainder considers only the first of these cases. Using this in the next innermost sum gives

$$\sum_{b=0}^{\infty}\left(\left(\frac{-\alpha\delta^{b}}{\alpha}\right)^{2}\sigma_{e}^{2}+\left(\frac{-\rho^{b}}{\alpha}\right)^{2}\sigma_{\varepsilon}^{2}\right) \tag{5.19}$$

Given that

$$\begin{aligned}\left|\delta\right|<1,\ lim_{b\to\infty}\delta^{2b}=0\\ \left|\rho\right|<1,\ lim_{b\to\infty}\rho^{2b}=0\end{aligned} \tag{5.20}$$

it follows that

$$\sum_{b=0}^{\infty}\left(\left(\frac{-\alpha\delta^{b}}{\alpha}\right)^{2}\sigma_{e}^{2}+\left(\frac{-\rho^{b}}{\alpha}\right)^{2}\sigma_{\varepsilon}^{2}\right)=\frac{\sigma_{e}^{2}}{1-\delta^{2}}+\frac{\sigma_{\varepsilon}^{2}}{\alpha^{2}(1-\rho^{2})} \tag{5.21}$$

The next step is to consider the probability weight term, which can now be expressed

[2] Here the terms within the sum are also sums, but as all of the terms within these are positive the inequality condition will be satisfied when the limits of these sums are taken.

as [3]

$$\lambda_i \sum_{z=0}^{\infty} (1-\lambda_i)^{n-z} \left(\frac{\sigma_e^2}{1-\delta^2} + \frac{\sigma_\varepsilon^2}{\alpha^2(1-\rho^2)} \right)$$
$$= \left(\frac{\sigma_e^2}{1-\delta^2} + \frac{\sigma_\varepsilon^2}{\alpha^2(1-\rho^2)} \right) \quad (5.22)$$

Lastly, given $\left|\beta\right| < 1$

$$\sum_{n=0}^{\infty} \beta^n \left(\frac{\sigma_e^2}{1-\delta^2} + \frac{\sigma_\varepsilon^2}{\alpha^2(1-\rho^2)} \right) \quad (5.23)$$
$$= \frac{1}{1-\beta} \left(\frac{\sigma_e^2}{1-\delta^2} + \frac{\sigma_\varepsilon^2}{\alpha^2(1-\rho^2)} \right) < \infty$$

Similarly the total cost must be convergent given the above condition on the discount factor and finite C, and the t + 1 loss is clearly finite also for meaningful parameter values, so the total value of the loss function is bounded below infinity, with the implication that for any N there will be some parameterization such that the optimal λ is interior.

For the asymptotic information distribution loss function, convergence follows by implication of equations 5.18, 5.21, 5.22 and 5.23.

A.2 Appendix to chapter 3

A.2.1 Model Specification

The model used in this chapter is principally that of Mankiw and Reis (2006b), with some differences which are highlighted as they are introduced. In turn, the model of Mankiw and Reis (2006b) is based on that of Erceg et al. (2000), in that it features monopolistically competitive price- and wage-setters. However, the models differ in

[3]This will only arise in the actual sum when n = ∞, so equivalently suppose the sum is over powers of the first bracket ranging from zero to one, giving the result shown.

that the Mankiw and Reis (2006b) model excludes capital and money, and imposes the assumption of sticky information. Hence the model is quite similar to the canonical simple New Keynesian model with sticky prices and wages, but with the Calvo-type price updating replaced by irregular updating of price plans[4]. The agents in the model are households, firms, government and a monetary policy authority.

Households consist of consumers, shoppers and workers, with each constituent type potentially possessing different information. The shopper is responsible for choosing the quantities of individual goods within the household's consumption and it is assumed that the shopper is always possessed of current price information, such that all individual prices are known and the composition of the consumption basket is identical across households. The consumer chooses the overall quantity of consumption while the worker chooses the nominal wage at which labour will be supplied and then provides such quantity as is required to satisfy the labour demand of firms. Both the consumer and worker receive new information probabilistically. Mankiw and Reis (2006b) allow the probability of arrival of new information in any period to differ between these two subtypes of agent. However, they note that the estimation results do not allow rejection of the hypothesis that the information acquisition probability is the same for both[5]. Therefore the following imposes a common probability for workers and consumers, with the consequence that the solution approach is slightly different, although this does not affect the eventual solution of the model.

Firms consist of a hiring department and a production/marketing department, where again it is possible for these two departments to possess different information from one another. The hiring department always possesses current nominal wage information and chooses the quantities of labour to hire from individual households. Hence the composite unit of labour used in production has the same composition across firms. As firms are monopolistically competitive the production/marketing department chooses the price at which to sell and then satisfies the consumer demand arising from this, using whatever quantity of composite labour is necessary to do so. Information is received

[4]Some of the steps of the following derivations which are common to both models are as in Gali (2008), chapter 6.

[5]Mankiw and Reis (2007)p611

probabilistically by the production/marketing department.

A.2.1.2 Shopper

Beginning with the problem of a shopper, if the composite consumption unit is defined as

$$C_{t,h} = \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1}} \tag{5.24}$$

for t the representative time period, h the index over the unit measure of households[6], f the index over the unit measure of firms and ν a stochastic preference parameter (the process for which is given below), then the shopper's problem is one of cost minimization

$$\min_{C_{t,h}(f)} \quad \int_0^1 P_t(f) C_{t,h}(f)\, df \;\; s.t.\; eq.5.24 \tag{5.25}$$

which has Lagrangian

$$\mathcal{L} = \int_0^1 P_t(f) C_{t,h}(f)\, df + \lambda_t \left\{ \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1}} - C_{t,h} \right\} \tag{5.26}$$

[6]Throughout, the notational convention is that a subscript index relates a variable to the principal agent within that context, while an index in parentheses relates a variable to a secondary agent. For example, in the shopper problem the price of good of type f is denoted $P_t(f)$ while in the firm's price-setting problem, the price set by firm f (which produces goods of type f) is denoted $P_{t,f}$. The index f will always refer to a sub-agent within a firm and the index h will always refer to a sub-agent within a household.

The first-order conditions and solutions to these are

$$\frac{\partial \mathcal{L}}{\partial \lambda_t} = \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1}} - C_{t,h} = 0$$

$$\rightarrow C_{t,h} = \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1}} \tag{5.27}$$

$$\frac{\partial \mathcal{L}}{\partial C_{t,h}(f)} = P_t(f) + \lambda_t \frac{\nu_t}{\nu_t - 1} \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1} - 1} \cdot \frac{\nu_t - 1}{\nu_t} C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t} - 1} = 0$$

$$\rightarrow P_t(f) + \lambda_t \left(\int_0^1 C_{t,h}\left(f\right)^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{1}{\nu_t - 1}} C_{t,h}\left(f\right)^{\frac{-1}{\nu_t}} = 0$$

$$\therefore C_{t,h}\left(f\right)^{\frac{-1}{\nu_t}} = \frac{P_t(f)}{\lambda_t C_{t,h}^{\frac{1}{\nu_t}}} \tag{5.28}$$

Then if two different varieties of the consumption good are considered

$$\frac{C_{t,h}\left(f_1\right)^{\frac{-1}{\nu_t}}}{C_{t,h}\left(f_2\right)^{\frac{-1}{\nu_t}}} = \frac{P_t(f_1)\lambda_t C_{t,h}^{\frac{1}{\nu_t}}}{P_t(f_2)\lambda_t C_{t,h}^{\frac{1}{\nu_t}}}$$

$$\rightarrow C_{t,h}\left(f_1\right) = \left(\frac{P_t(f_1)}{P_t(f_2)} \right)^{-\nu_t} C_{t,h}\left(f_2\right) \tag{5.29}$$

The price index is defined as

$$P_t = \left(\int_0^1 P_t(f)^{1-\nu_t} df \right)^{\frac{1}{1-\nu_t}} \tag{5.30}$$

and total household expenditures are

$$Z_{t,h} = \int_0^1 P_t(f) C_{t,h}(f) df \tag{5.31}$$

Therefore, using equation 5.29

$$\begin{aligned}
Z_{t,h} &= \int_0^1 P_t(f_1) \left(\frac{P_t(f_1)}{P_t(f_2)} \right)^{-\nu_t} C_{t,h}(f_2) \, df_1 \\
&= \frac{C_{t,h}(f_2)}{P_t(f_2)^{-\nu_t}} \int_0^1 P_t(f_1)^{1-\nu_t} df_1 \\
&= C_{t,h}(f_2) \, P_t(f_2)^{\nu_t} P_t^{1-\nu_t} \\
&= C_{t,h}(f) \, P_t(f)^{\nu_t} P_t^{1-\nu_t}
\end{aligned} \tag{5.32}$$

by definition of the price index (equation (5.30)) and for f the general index over firms. Then equation 5.32 can be re-arranged to

$$\begin{aligned}
Z_{t,h} &= C_{t,h}(f) \, P_t(f)^{\nu_t} P_t^{1-\nu_t} \\
\frac{Z_{t,h}}{P_t} \frac{1}{P_t^{-\nu_t}} &= C_{t,h}(f) \, P_t(f)^{\nu_t} \\
\therefore C_{t,h}(f) &= \left(\frac{P_t(f)}{P_t} \right)^{-\nu_t} \frac{Z_{t,h}}{P_t}
\end{aligned} \tag{5.33}$$

which is that the demand of household h for good variety f depends upon the price of the variety, the price level and the household's total real consumption demand. Given the definition of composite consumption in equation 5.24, equation 5.33 can be substituted therein

$$C_{t,h} = \left(\int_0^1 \left\{ \left(\frac{P_t(f)}{P_t} \right)^{-\nu_t} \frac{Z_{t,h}}{P_t} \right\}^{\frac{\nu_t - 1}{\nu_t}} df \right)^{\frac{\nu_t}{\nu_t - 1}}$$

$$= Z_{t,h} P_t^{\nu_t - 1} \left(\int_0^1 P_t(f)^{1-\nu_t} df \right)^{\frac{\nu_t}{\nu_t - 1}}$$

$$= Z_{t,h} P_t^{\nu_t - 1} P_t^{-\nu_t} = Z_{t,h} P_t^{-1} \therefore Z_{t,h} = C_{t,h} P_t \tag{5.34}$$

Equation 5.34 shows that the household consumption expenditure across all varieties of goods can be expressed as the product of household total consumption and the price level. As this derives from the solution of the shopper's problem and all shoppers are fully informed this will hold for all households, given their decisions about total consumption. Hence the right-hand side of equation 5.34 can be used to represent consumption expenditures in the generic household budget constraint.

A.2.1.3 Hiring Department

The other fully informed agent sub-type is the hiring department of the firm, which is tasked with minimizing the cost of hiring a composite unit of labour, where labour input varieties are aggregated according to

$$L_{t,f} = \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} \tag{5.35}$$

As previously, f indexes firms and h indexes households. $L_{t,f}$ is the total labour input used by firm f in period t and γ is a stochastic elasticity of substitution (the process for which is specified below). Given that total wage expenditures are $\int_0^1 W_t(h) L_{t,f}(h)\, dh$

the relevant Lagrangian is

$$\mathcal{L} = \int_0^1 W_t(h) L_{t,f}(h)\, dh + \lambda_t \left\{ \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} - N_{t,f} \right\} \tag{5.36}$$

The first-order conditions and solutions to these are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda_t} &= \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} - L_{t,f} = 0 \\ \rightarrow L_{t,f} &= \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} \end{aligned} \tag{5.37}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial L_{t,f}(h)} = W_t(h) + \lambda_t \frac{\gamma_t}{\gamma_t - 1} \left(\int_0^1 L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1} - 1} & \\ \cdot \frac{\gamma_t - 1}{\gamma_t} L_{t,f}(h)^{\frac{\gamma_t - 1}{\gamma_t} - 1} &= 0 \\ \rightarrow W_t(h) + \lambda_t L_{t,f}^{\frac{1}{\gamma_t}} L_{t,f}(h)^{\frac{-1}{\gamma_t}} &= 0 \\ \therefore L_{t,f}(h) &= \frac{-W_t(h)}{\lambda_t L_{t,f}^{\frac{1}{\gamma_t}}} \end{aligned} \tag{5.38}$$

If two different varieties of labour are considered

$$\begin{aligned} \frac{L_{t,f}\left(h_1\right)^{\frac{-1}{\gamma_t}}}{L_{t,f}\left(h_2\right)^{\frac{-1}{\gamma_t}}} &= \frac{-W_t(h_1) \lambda_t L_{t,f}^{\frac{1}{\gamma_t}}}{-W_t(h_2) \lambda_t L_{t,f}^{\frac{1}{\gamma_t}}} \\ \therefore L_{t,f}(h_1) &= \left(\frac{W_t(h_1)}{W_t(h_2)} \right)^{-\gamma_t} L_{t,f}(h_2) \end{aligned} \tag{5.39}$$

Total wage costs for firm f are $X_{t,f} = \int_0^1 W_t(h) L_{t,f}(h) dh$ and the wage index may be defined as

$$W_t = \left(\int_0^1 W_t(h)^{1-\gamma_t} dh \right)^{\frac{1}{1-\gamma_t}} \tag{5.40}$$

Combining these two gives

$$\begin{aligned} X_{t,f} &= \int_0^1 W_t(h_1) \left(\frac{W_t(h_1)}{W_t(h_2)} \right)^{-\gamma_t} L_{t,f}(h_2)\, dh_1 \\ &= L_{t,f}(h_2) W_t(h_2)^{\gamma_t} \int_0^1 W_t(h_1)^{1-\gamma_t}\, dh_1 \\ &= L_{t,f}(h_2) W_t(h_2)^{\gamma_t} W_t^{1-\gamma_t} \end{aligned} \tag{5.41}$$

Using equations 5.41 and 5.35

$$\begin{aligned} L_{t,f}(h) &= \left(\frac{W_t(h)}{W_t} \right)^{-\gamma_t} \frac{X_{t,f}}{W_t} \\ L_{t,f} &= \left(\int_0^1 \left\{ \left(\frac{W_t(h)}{W_t} \right)^{-\gamma_t} \frac{X_{t,f}}{W_t} \right\}^{\frac{\gamma_t - 1}{\gamma_t}} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} \\ &= X_{t,f} W_t^{\gamma_t - 1} \left(\int_0^1 W_t(h)^{1-\gamma_t} dh \right)^{\frac{\gamma_t}{\gamma_t - 1}} \\ &= X_{t,f} W_t^{\gamma_t - 1} W_t^{-\gamma_t} \\ \therefore X_{t,f} &= L_{t,f} W_t \end{aligned} \tag{5.42}$$

which shows that total labour costs for firm f in period t can be expressed as the product of the composite labour input and the aggregate wage level. As all hiring departments are fully informed by assumption this will apply to all firms and the right-hand side of equation 5.42 can be used in the generic firm profit function.

A.2.1.4 Consumer/worker

The household's preferences over consumption and leisure are represented by the period utility function

$$U(C_{t,h},\ L_{t,h}) = ln\Big(C_{t,h}\Big) - \frac{L_{t,h}^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \tag{5.43}$$

where $L_{t,h}$ is the quantity of labour supplied by household h in period t. It is assumed that all households have the same preferences and hence the same utility function. The budget constraint is

$$P_t C_{t,h} + \frac{B_{t,h}}{I_t} \leq B_{t-1,h} + W_{t,h} L_{t,h} + T_t \tag{5.44}$$

where $B_{t,h}$ is the nominal redemption value of one-period bonds purchased by household h in period t and I_t is the gross nominal interest rate applicable to period t, such that the current cost of purchasing the future bond payment is given by the ratio between B and I. T_t represents net lump sum transfers to and from the government, plus transfers from an insurance contract which ensures that all households have the same wealth as each other at the beginning of each period.

The decision variables for the household are $C_{t,h}, W_{t,h}$ and $B_{t,h}$. As a wage-setting monopolist the household supplies however much labour is demanded by firms at the specified wage. Equation 5.42 implicitly defines the demand of firm f for labour of type h as

$$L_{t,f}(h) = \left(\frac{W_t(h)}{W_t}\right)^{-\gamma_t} \frac{X_{t,f}}{W_t} = \left(\frac{W_t(h)}{W_t}\right)^{-\gamma_t} N_{t,f} \tag{5.45}$$

so the total demand for labour of type h is

$$\begin{aligned}
\int_0^1 L_{t,f}(h)\,df &= \int_0^1 \left(\frac{W_t(h)}{W_t}\right)^{-\gamma_t} L_{t,f}\,df \\
&= \left(\frac{W_t(h)}{W_t}\right)^{-\gamma_t} \int_0^1 L_{t,f}\,df \\
\therefore L_t(h) &= \left(\frac{W_t(h)}{W_t}\right)^{-\gamma_t} L_t
\end{aligned} \tag{5.46}$$

where L_t is aggregate labour demand in the economy, which by definition is equal to the term which it replaces in equation 5.46. As the market-clearing quantity of labour is determined by demand, $L_{t,h} = L_t(h)$. At the beginning of any period t the optimization problem for the consumer/worker, given that the information possessed is k periods old, is to choose $C_{t,h}, W_{t,h}$ and $B_{t,h}$ in order to maximize the expected discounted utility

$$max.\, E_{t-k}\left[\sum_{s=0}^{\infty} \xi^s \left(ln\Big(C_{t+s,h}\Big) - \frac{L_{t+s,h}^{1+\frac{1}{\psi}} + 1}{1+\frac{1}{\psi}}\right)\right]$$

subject to

$$P_{t+s}C_{t+s,h} + \frac{B_{t+s,h}}{I_{t+s}} \leq B_{t+s-1,h} + W_{t+s,h}L_{t+s,h} + T_{t+s} \tag{5.47}$$

$$L_{t+s,h} = \left(\frac{W_{t+s,h}}{W_{t+s}}\right)^{-\gamma_{t+s}} L_{t+s} \tag{5.48}$$

$$0 = \mu_{t+s}C_{t+s,h} \tag{5.49}$$

$$0 = \varpi_{t+s}W_{t+s,h} \tag{5.50}$$

$$lim_{s\to\infty}E_t\Big[B_{t+s,h}\Big] \geq 0\,\forall t \tag{5.51}$$

Equations 5.47 and 5.48 are the budget constraint and labour demand constraint, respectively, discussed above. Equations 5.49 and 5.50 are non-negativity contraints for consumption and the wage, respectively. Equation 5.51 is a no-Ponzi scheme condition. As in Mankiw and Reis (2006b) it is assumed that the bond market clears with zero net supply. There is no non-negativity constraint relating to hours because the utility function is such that the solution for hours should be interior and stationary[7]. Furthermore, the marginal utility of consumption is such that $lim_{c\to 0}U_c(.) \to \infty$, so consumption will be strictly positive and $\mu = 0$ always. Similarly, a zero wage would imply arbitrarily large/ill-defined labour demand (given that hours are bounded), so a strictly positive wage will imply that $\varpi = 0$ always. Consequently constraints 5.49 and 5.50 are guaranteed to be satisfied and so need not be explicitly included below. The Lagrangian for this problem, given information of age k, is[8]

$$\begin{aligned}\mathcal{L}_t = E_{t-k}\Bigg[\sum_{s=0}^{\infty}\xi^s\Bigg\{ln\Big(C_{t+s,h}\Big) - \frac{\left(\left(\frac{W_{t+s,h}}{W_{t+s}}\right)^{-\gamma_{t+s}}L_{t+s}\right)^{1+\frac{1}{\psi}}+1}{1+\frac{1}{\psi}} \\ +\lambda_{t+s}(B_{t+s-1,h} + W_{t+s,h}\Big(\frac{W_{t+s,h}}{W_{t+s}}\Big)^{-\gamma_{t+s}}L_{t+s} \\ +T_{t+s} - P_{t+s} - \frac{B_{t+s,h}}{I_{t+s}})\Bigg\}\Bigg] \qquad (5.52)\end{aligned}$$

[7]As discussed, for example, by Heer and Maussner (2009), chapter 1.

[8]In the remaining equations of this subsection λ denotes a Lagrange multiplier rather than a probability.

The first-order conditions are[9]

$$\frac{\partial \mathcal{L}_t}{\partial C_{t,h}} = E_{t-k}\left[\xi^0\left(\frac{1}{C_{t,h}} - \lambda_t P_t\right)\right] = 0 \tag{5.53}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_t}{\partial W_{t,h}} = E_{t-k}\Bigg[\xi^0\Bigg(\gamma_t\left(\left(\frac{W_{t,h}}{W_t}\right)^{-\gamma_t} L_t\right)^{1+\frac{1}{\psi}} W_{t,h}^{-1}\Bigg) \\ +\lambda_t(1-\gamma_t)\left(\frac{W_{t,h}}{W_t}\right)^{-\gamma_t} L_t\Bigg] = 0\end{aligned} \tag{5.54}$$

$$\frac{\partial \mathcal{L}}{\partial B_{t,h}} = E_{t-k}\left[-\xi^0\frac{\lambda_t}{I_t} + \xi^1\lambda_{t+1}\right] = 0 \tag{5.55}$$

$$\frac{\partial \mathcal{L}_t}{\partial \lambda_t} = B_{t-1,h} + W_{t,h}\left(\frac{W_{t,h}}{W_t}\right)^{-\gamma_t} L_t + T_t - P_t - \frac{B_{t,h}}{I_t} = 0 \tag{5.56}$$

Assuming where necessary that covariances between different variables are negligible, these can be used to derive the following:

Using equations 5.53 and 5.55

$$\begin{aligned}\lambda_t = E_{t-k}\left[\frac{1}{C_{t,h}P_t}\right] \rightarrow \lambda_{t+1} &= E_{t-k}\left[\frac{1}{C_{t+1,h}P_{t+1}}\right] \\ \therefore \frac{\lambda_{t+1}}{\lambda_t} &= E_{t-k}\left[\frac{C_{t,h}P_t}{C_{t+1,h}P_{t+1}}\right] \\ \frac{\xi\lambda_{t+1}}{\lambda_t} &= E_{t-k}\left[\frac{1}{I_t}\right] \\ \xi E_{t-k}\left[\frac{C_{t,h}P_t}{C_{t+1,h}P_{t+1}}\right] &= E_{t-k}\left[\frac{1}{I_t}\right] \\ E_{t-k}\Big[C_{t,h}\Big] \equiv C_{t,h}^k &= \frac{1}{\xi}E_{t-k}\left[\frac{C_{t+1,h}P_{t+1}}{I_tP_t}\right]\end{aligned} \tag{5.57}$$

[9]Here the derivatives are taken with respect to variables dated $t+s$. The equations shown use $s=0$ to ease the notation, but they hold for all $s>0$ also. ξ^0 is used rather than 1 in order that the general case can be discerned more easily.

[10]Using equation 5.54 and the first line of the previous equations

$$\begin{aligned} E_{t-k}\left[W_{t,h}\right] \equiv W_{t,h}^{k} &= E_{t-k}\left[\frac{\gamma_t}{\lambda_t(\gamma_t-1)}\left(\left(\frac{W_{t,h}}{W_t}\right)^{-\gamma_t} L_t\right)^{\frac{1}{\psi}}\right] \\ &= E_{t-k}\left[\frac{\gamma_t}{\lambda_t(\gamma_t-1)} L_{t,h}^{\frac{1}{\psi}}\right] \\ &= E_{t-k}\left[\frac{\gamma_t}{\gamma_t-1} C_{t,h} P_t L_{t,h}^{\frac{1}{\psi}}\right] \end{aligned} \tag{5.58}$$

A.2.1.5 Government

Before considering the profit-maximization behaviour of firms it is necessary to introduce the government sector. The representation of fiscal policy within the model is simply to suppose that there is a stochastic government demand for goods, such that government consumes the proportion $1-\frac{1}{G_t}$ of total output in period t, where the stochastic process for G_t is given below. Consequently households consume the proportion $\frac{1}{G_t}$ of output of each good

$$C_t(f) = \frac{1}{G_t} Y_{t,f} \rightarrow Y_{t,f} = G_t C_t(f) \tag{5.59}$$

and aggregate household demand for good f, using equation 5.33, is given by

$$\begin{aligned} C_t(f) &= \int_0^1 C_{t,h}(f)\,dh = \int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} C_{t,h}\,dh \\ &= \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} \int_0^1 C_{t,h}\,dh = \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} C_t \end{aligned} \tag{5.60}$$

[10]In this notation where superscript k denotes the age of the information available, the absence of a superscript implies k = 0, i.e. current information is available.

Therefore the total demand for good f, given the realization of the government demand shock, is

$$Y_{t,f} = G_t \left(\frac{P_t(f)}{P_t} \right)^{-\nu_t} C_t \tag{5.61}$$

A.2.1.6 Pricing/Production Department

Within firms the pricing department is tasked with setting the price at which output is sold and choosing the quantity of composite labour to hire in order to maximize the profit earned. Real profits are given by

$$\pi_{t,f} = \frac{P_{t,f} Y_{t,f}}{P_t} - \frac{W_t L_{t,f}}{P_t} \tag{5.62}$$

where $Y_{t,f}$ is the output sold by firm f in period t and equation 5.42 has been used to express the total labour costs. The constraints on this optimization are the demand constraint (equation 5.61) and the production function

$$Y_{t,f} = A_t L_{t,f}^{\beta} \tag{5.63}$$

where A_t is a stochastic technology parameter that is common to all firms (the process for which is given below). $\beta \in (0, 1)$ is fixed and represents the degree of diminishing marginal returns to labour. There is no productive capital in this model and so the decisions made in each period do not have any direct intertemporal consequences. Therefore the pricing/production department has as its objective single period profits as in equation 5.62. Information is received probabilistically, so the decisions may be based on old information. Assuming strictly positive price and labour input, this gives the Lagrangian for a firm with information of age j as[11]

$$\mathcal{L}_t = E_{t-j}\left[\left(\frac{P_{t,f}}{P_t}\right)^{1-\nu_t} C_t G_t - \frac{W_t L_{t,f}}{P_t} \right.$$
$$\left. +\iota_t\left\{A_t L_{t,f}^{\beta} - \left(\frac{P_{t,f}}{P_t}\right)^{-\nu_t} C_t G_t\right\}\right] \tag{5.64}$$

The first-order conditions are

$$\frac{\partial \mathcal{L}_t}{\partial P_{t,f}} = E_{t-j}\left[(1-\nu_t)C_t G_t \left(\frac{P_{t,f}}{P_t}\right)^{1-\nu_t} P_{t,f}^{-1} \right.$$
$$\left. +\iota_t \nu_t C_t G_t \left(\frac{P_{t,f}}{P_t}\right)^{-\nu_t} P_{t,f}\right] = 0 \tag{5.65}$$

$$\frac{\partial \mathcal{L}_t}{\partial L_{t,f}} = E_{t-j}\left[-\frac{W_t}{P_t} + \iota_t \beta A_t L_{t,f}^{\beta-1}\right] = 0 \tag{5.66}$$

Under the same assumptions as for the re-arrangement of the consumer/worker first-order conditions, the above lead to

$$\iota_t = E_{t-j}\left[\frac{W_t L_{t,f}}{P_t \beta A_t L_{t,f}^{\beta}}\right] = E_{t-j}\left[\frac{W_t L_{t,f}}{P_t \beta Y_{t,f}}\right]$$
$$E_{t-j}\left[\frac{P_{t,f}}{P_t} Y_{t,f} P_{t,f}^{-1}\right] = E_{t-j}\left[\frac{\nu_t}{\nu_t - 1} Y_{t,f} P_{t,f}^{-1} \iota_t\right]$$
$$\therefore E_{t-j}\left[P_{t,f}\right] \equiv P_{t,f}^{j} = E_{t-j}\left[\frac{\nu_t W_t L_{t,f}}{(\nu_t - 1)\beta Y_{t,f}}\right] \tag{5.67}$$

The next step of the model solution is to derive log-linear approximations.

[11] The single constraint term here comes from imposing demand-determined production, i.e. combining equations 5.61 and 5.63.

A.2.2 Log-Linearized Model

Here the approach detailed by McCandless (2008) (see chapter 6) is used, specifically each variable X_t is replaced with $\bar{X}e^{\tilde{x}_t}$, where $\tilde{x}_t \equiv ln\left(X_t\right) - ln\left(\bar{X}\right)$ and $\bar{X}$ is the value of the variable in the deterministic steady state. After this substitution the exponential terms are approximated by first-order Taylor series, so the resulting equations are linear in the log deviations from steady state.

Starting with equation 5.57

$$C^k_{t,h} = \frac{1}{\xi} E_{t-k}\left[\frac{C_{t+1,h}P_{t+1}}{I_t P_t}\right]$$

$$\bar{C}e^{\tilde{c}^k_{t,h}} = \frac{1}{\xi}\frac{\bar{C}\bar{P}}{\bar{I}\bar{P}} E_{t-k}\left[e^{\tilde{c}_{t+1,h}+\tilde{p}_{t+1}-\tilde{p}_t-\tilde{i}_t}\right] \tag{5.68}$$

In the steady state $\bar{C} = \frac{1}{\xi}\frac{\bar{C}}{\bar{I}} \Rightarrow \bar{I} = \frac{1}{\xi}$ and

$$e^{\tilde{c}^k_{t,h}} = E_{t-k}\left[e^{\tilde{c}_{t+1,h}+\tilde{p}_{t+1}-\tilde{p}_t-\tilde{i}_t}\right]$$

$$\therefore \tilde{c}^k_{t,h} \approx E_{t-k}\left[\tilde{c}_{t+1,h} + \tilde{p}_{t+1} - \tilde{p}_t - \tilde{i}_t\right] \tag{5.69}$$

Next, consider equation 5.58

$$W^k_{t,h} = E_{t-k}\left[\frac{\gamma_t}{\gamma_t - 1} C_{t,h} P_t L^{\frac{1}{\psi}}_{t,h}\right]$$

$$\bar{W}e^{\tilde{w}^k_{t,h}} = E_{t-k}\left[\frac{\bar{\gamma}}{\bar{\gamma}-1}\bar{C}\bar{P}\bar{L}^{\frac{1}{\psi}} e^{\tilde{\gamma}_t - \widetilde{\gamma_t - 1} + \tilde{c}_{t,h} + \tilde{p}_t + \frac{1}{\psi}\tilde{l}_{t,h}}\right] \tag{5.70}$$

In the steady state $\bar{W} = \frac{\bar{\gamma}}{\bar{\gamma}-1}\bar{C}\bar{P}\bar{L}^{\frac{1}{\psi}}$ and

$$e^{\tilde{w}^k_{t,h}} = E_{t-k}\left[e^{\tilde{\gamma}_t - \widetilde{\gamma_t - 1} + \tilde{c}_{t,h} + \tilde{p}_t + \frac{1}{\psi}\tilde{l}_{t,h}}\right]$$

$$\therefore \tilde{w}^k_{t,h} \approx E_{t-k}\left[\tilde{\gamma}_t - \widetilde{\gamma_t - 1} + \tilde{c}_{t,h} + \tilde{p}_t + \frac{1}{\psi}\tilde{l}_{t,h}\right] \tag{5.71}$$

Given that

$$\begin{aligned}
\widetilde{\gamma_t - 1} = ln(\gamma_t - 1) - ln(\overline{\gamma} - 1) &= ln\left(\frac{\gamma_t - 1}{\bar{\gamma} - 1}\right) \\
&= ln\left(\frac{\overline{\gamma} - 1}{\overline{\gamma} - 1} + \frac{\gamma_t - 1 - (\overline{\gamma} - 1)}{\overline{\gamma} - 1}\right) \\
&= ln\left(1 + \frac{\gamma_t - \bar{\gamma}}{\bar{\gamma} - 1}\right) \approx \frac{\gamma_t - \bar{\gamma}}{\bar{\gamma} - 1} \quad \rightarrow \\
\widetilde{\gamma_t - 1}(\bar{\gamma} - 1) &\approx \gamma_t - \bar{\gamma} \\
\tilde{\gamma}_t \approx \frac{\gamma_t - \bar{\gamma}}{\bar{\gamma}} \rightarrow \gamma_t - \bar{\gamma} \approx \bar{\gamma}\tilde{\gamma}_t &\approx \widetilde{\gamma_t - 1}(\bar{\gamma} - 1) \\
\widetilde{\gamma_t - 1} &= \frac{\bar{\gamma}}{\bar{\gamma} - 1}\tilde{\gamma}_t
\end{aligned} \tag{5.72}$$

Therefore equation 5.71 can be written as

$$\tilde{w}^k_{t,h} \approx E_{t-k}\left[\frac{-\tilde{\gamma}_t}{\bar{\gamma} - 1} + \tilde{c}_{t,h} + \tilde{p}_t + \frac{\tilde{l}_{t,h}}{\psi}\right] \tag{5.73}$$

From equation 5.67

$$P^j_{t,f} = E_{t-j}\left[\frac{\nu_t W_t L_{t,f}}{(\nu_t - 1)\beta Y_{t,f}}\right]$$

$$\bar{P}e^{\tilde{p}^j_t} = E_{t-j}\left[\frac{\bar{\nu}}{\bar{\nu} - 1}\frac{\bar{W}\bar{L}}{\beta\bar{Y}}e^{\tilde{\nu}_t - \widetilde{\nu_t - 1} + \tilde{w}_t + \tilde{l}_{t,f} - \tilde{y}_{t,f}}\right] \tag{5.74}$$

In the steady state $\bar{P} = \frac{\bar{\nu}}{\bar{\nu}-1}\frac{\bar{W}\bar{L}}{\beta\bar{Y}}$, so

$$\begin{aligned} e^{\tilde{p}_t^j} &= E_{t-j}\left[e^{\tilde{\nu}_t - \widetilde{\nu_t - 1} + \tilde{w}_t + \tilde{l}_{t,f} - \tilde{y}_{t,f}}\right] \\ \tilde{p}_t^j &= E_{t-j}\left[\tilde{\nu}_t - \widetilde{\nu_t - 1} + \tilde{w}_t + \tilde{l}_{t,f} - \tilde{y}_{t,f}\right] \end{aligned} \tag{5.75}$$

Then using a re-arrangement similar to that in equation 5.72 this can be expressed as

$$\tilde{p}_t^j = E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t + \tilde{l}_{t,f} - \tilde{y}_{t,f}\right] \tag{5.76}$$

Next log-linearize the production function (equation 5.63)

$$\begin{aligned} Y_{t,f} &= A_t L_{t,f}^{\beta} \\ \bar{Y}e^{\tilde{y}_{t,f}} &= \bar{A}\bar{L}^{\beta}e^{\tilde{a}_t + \beta\tilde{n}_{t,f}} \\ \bar{Y} &= \bar{A}\bar{L}^{\beta} \\ e^{\tilde{y}_{t,f}} &= e^{\tilde{a}_t + \beta\tilde{l}_{t,f}} \\ \tilde{y}_{t,f} &\approx \tilde{a}_t + \beta\tilde{l}_{t,f} \end{aligned} \tag{5.77}$$

The market clearing condition for good f is $Y_{t,f} = G_t\int_0^1 C_{t,h}(f)\,dh$ so the aggregate market clearing condition is[12]

$$Y_t = \left(\int_0^1 Y_{t,f}^{\frac{\nu_t-1}{\nu_t}}\,df\right)^{\frac{\nu_t}{\nu_t-1}} = \left(\int_0^1 \left(G_t \int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} C_{t,h}\,dh\right)^{\frac{\nu_t-1}{\nu_t}} df\right)^{\frac{\nu_t}{\nu_t-1}}$$
$$= \left(\int_0^1 \left(G_t \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} C_t\right)^{\frac{\nu_t-1}{\nu_t}} df\right)^{\frac{\nu_t}{\nu_t-1}}$$
$$= G_t C_t \left(\int_0^1 \left(\left(\frac{P_t(f)}{P_t}\right)^{-\nu_t}\right)^{\frac{\nu_t-1}{\nu_t}} df\right)^{\frac{\nu_t}{\nu_t-1}} = G_t C_t \tag{5.78}$$

where the last line uses the definition of the price index, equation 5.30. This leads to the log-linear approximation

$$\bar{Y} e^{\tilde{y}_t} = \bar{G}\bar{C} e^{\tilde{g}_t + \tilde{c}_t}$$
$$\therefore \tilde{y}_t = \tilde{g}_t + \tilde{c}_t \tag{5.79}$$

The price level is given by equation 5.30. Given that all firms have identical technology and the disturbances are aggregate rather than idiosyncratic, all firms in period t with up-to-date information will choose the same price[13]. The probability of receiving new information for a firm is λ, so in period t the proportion of firms receiving new information will be λ and if these firms are grouped together index-wise the price level can be written as

$$P_t = \left(\int_0^1 P_t(f)^{1-\nu_t} df\right)^{\frac{1}{1-\nu_t}} = \left(\int_\lambda^1 P_t(f)^{1-\nu_t} df + \int_0^\lambda P_t^{(k=0)\,1-\nu_t} df\right)^{\frac{1}{1-\nu_t}}$$
$$= \left(\int_\lambda^1 P_t(f)^{1-\nu_t} df + \lambda P_t^{(k=0)\,1-\nu_t}\right)^{\frac{1}{1-\nu_t}} \tag{5.80}$$

[12] This implicitly uses equation 5.60.

[13] The following equations do not use the f subscript on prices because of this. Firm prices can be distinguished from the price level by the continued use of a superscript to denote the age of information upon which the firm price is based.

Similarly, the proportion $\lambda(1-\lambda)$ will have information that is one period old, $\lambda(1-\lambda)^2$ will have information that is two periods old and so on. Therefore the price level can be written

$$P_t^{1-\nu_t} = \lambda \sum_{k=0}^{\infty}(1-\lambda)^k P_t^{k\,1-\nu_t} \tag{5.81}$$

Log-linearizing the price level in this form

$$\begin{aligned}
\bar{P}^{1-\nu_t} e^{\tilde{p}_t(1-\nu_t)} &= \bar{P}^{1-\nu_t}\lambda \sum_{k=0}^{\infty}(1-\lambda)^k e^{\tilde{p}_t^k(1-\nu_t)} \\
e^{\tilde{p}_t(1-\nu_t)} &= \lambda \sum_{k=0}^{\infty}(1-\lambda)^k e^{\tilde{p}_t^k(1-\nu_t)} \\
(1-\bar{\nu})\tilde{p}_t &\approx \lambda \sum_{k=0}^{\infty}(1-\lambda)^k(1-\bar{\nu})\tilde{p}_t^k
\end{aligned}$$

$$\tilde{p}_t \approx \lambda \sum_{k=0}^{\infty}(1-\lambda)^k \tilde{p}_t^k \tag{5.82}$$

Aggregate consumption is defined as

$$C_t = \int_0^1 C_{t,h} dh \tag{5.83}$$

All households with information of the same age will choose the same level of consumption and households receive new information with probability ω in each period[14].

[14] Mankiw and Reis (2006b) have different update probabilities between consumers and workers, whereas here there is a single probability for the entire household. This gives a different specification of the optimization problem, however the same solution is still obtained, with the exception of there being fewer information probabilities. In their estimation, Mankiw and Reis find that the consumer and worker probabilities are sufficiently close that a hypothesis of equality cannot be rejected. Therefore it may be unlikely that the prior imposition of equal probabilities here will cause significant differences in the results.

Therefore, by the same reasoning as for the price level, the aggregate consumption will be

$$C_t = \omega \sum_{k=0}^{\infty} (1-\omega)^k C_t^k \tag{5.84}$$

which log-linearizes to

$$\begin{aligned} \bar{C} e^{\widetilde{c_t}} &= \omega \sum_{k=0}^{\infty} (1-\omega)^k \bar{C} e^{\tilde{c_t}^k} \\ e^{\widetilde{c_t}} &= \omega \sum_{k=0}^{\infty} (1-\omega)^k e^{\tilde{c_t}^k} \\ \widetilde{c_t} &= \omega \sum_{k=0}^{\infty} (1-\omega)^k \tilde{c_t}^k \end{aligned} \tag{5.85}$$

The wage level is defined as

$$W_t = \left(\int_0^1 W_t(h)^{1-\gamma_t} dh \right)^{\frac{1}{1-\gamma_t}} \tag{5.86}$$

All households with the same information will choose the same nominal wage, so, analogous to the price level and aggregate consumption, this can be expressed as

$$W_t^{1-\gamma_t} = \omega \sum_{k=0}^{\infty} (1-\omega)^k W_t^{k\,1-\gamma_t} \tag{5.87}$$

Log-linearizing this gives

$$\begin{aligned}
\bar{W}^{1-\gamma_t} e^{\tilde{w}_t(1-\gamma_t)} &= \bar{W}^{1-\gamma_t} \omega \sum_{k=0}^{\infty} (1-\omega)^k e^{\tilde{w}_t^k (1-\gamma_t)} \\
e^{\tilde{w}_t(1-\gamma_t)} \approx e^{0(1-\bar{\gamma})} + (1-\bar{\gamma}) e^{0(1-\bar{\gamma})}(\tilde{w}_t - 0) + 0 e^{0(1-\bar{\gamma})}&(1-\gamma_t-(1-\bar{\gamma})) \\
&= 1 + (1-\bar{\gamma})\tilde{w}_t \\
\tilde{w}_t &= \omega \sum_{k=0}^{\infty} (1-\omega)^k \tilde{w}_t^k \qquad (5.88)
\end{aligned}$$

Monetary policy is assumed to follow the rule

$$E_t\left[R_t \frac{P_{t+1}}{P_t}\right] \equiv I_t = \bar{I} \left(\frac{Y_t}{Y_t^n}\right)^{\phi_y} \left(\frac{P_{t+1}}{P_t}\right)^{\phi_p} \varepsilon_t^{-1} \qquad (5.89)$$

where the first equality defines the nominal interest rate. Y_t^n is the natural level of output, where here natural refers to that which would obtain if all agents had up-to-date information (rather than referring to a state without monopolistic competition).

The log-linear relation is

$$\begin{aligned}
\bar{I} e^{\tilde{i}_t} &= \frac{\overline{RP}}{\bar{P}} e^{\tilde{r}_t + E_t[\tilde{p}_{t+1} - \tilde{p}_t]} \\
\therefore \tilde{i}_t &= \tilde{r}_t + E_t\left[\tilde{p}_{t+1} - \tilde{p}_t\right] \qquad (5.90)
\end{aligned}$$

and the log-linear policy rule is[15]

$$\tilde{i}_t = \phi_y(\tilde{y}_t - \tilde{y}_t^n) + \phi_p(\tilde{p}_t - \tilde{p}_{t-1}) - \tilde{\varepsilon}_t \qquad (5.91)$$

Given equations 5.61 and 5.78 the market-clearing condition for a good f can be

[15] ε_t now refers to the log of the disturbance term in equation 5.89.

expressed as

$$Y_{t,f} = \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} G_t C_t$$
$$Y_{t,f} = \left(\frac{P_t(f)}{P_t}\right)^{-\nu_t} Y_t \tag{5.92}$$

Log-linearization of this condition gives

$$y_{t,f} = -\nu_t(p_{t,f} - p_t) + y_t$$
$$\tilde{y}_{t,f} + \bar{y}_f = -(\tilde{\nu}_t + \bar{\nu})(\tilde{p}_{t,f} + \bar{p}_f - \tilde{p}_t - \bar{p}) + \tilde{y}_t + \bar{y} \tag{5.93}$$

With identical technologies $\bar{y}_f = \bar{y}$ and $\bar{p}_f = \bar{p}$. Using also the approximation that products of log deviation terms are zero, the above becomes

$$\tilde{y}_{t,f} = -\bar{\nu}(\tilde{p}_{t,f} - \tilde{p}_t) + \tilde{y}_t \tag{5.94}$$

Equation 5.46 gives the overall demand for labour of type h. Given that labour supply adjusts to satisfy demand, this implies

$$L_{t,h} = \left(\frac{W_{t,h}}{W_t}\right)^{-\gamma_t} L_t \tag{5.95}$$

Then taking logarithms

$$l_{t,h} = -\gamma_t(w_{t,h} - w_t) + l_t \tag{5.96}$$

and then substituting the deviations and steady state values into this gives

$$\tilde{l}_{t,h} + \bar{l}_h = -(\tilde{\gamma}_t + \bar{\gamma})(\tilde{w}_{t,h} + \bar{w}_h - \tilde{w}_t - \bar{w}) + \tilde{l}_t + \bar{l}$$
$$\tilde{l}_{t,h} = -\bar{\gamma}(\tilde{w}_{t,h} - \tilde{w}_t) + \tilde{l}_t \tag{5.97}$$

The equations for household consumption, wages and firm price (equations 5.69, 5.73 and 5.76) contain individual variables, so in order to simplify the solution these variables are now substituted for or otherwise replaced. Starting with equation 5.69, solving forward gives

$$\begin{aligned}\tilde{c}^k_{t,h} &\approx E_{t-k}\left[\tilde{c}_{t+1,h} + \tilde{p}_{t+1} - \tilde{p}_t - \tilde{i}_t\right] \\ &= E_{t-k}\left[\tilde{c}_{t+2,h} + \Delta\tilde{p}_{t+2} - \tilde{i}_{t+1} + \Delta\tilde{p}_{t+1} - \tilde{i}_t\right] \\ &= E_{t-k}\left[\tilde{y}^n_\infty + \sum_{m=0}^{\infty} -(\tilde{i}_{t+m} - \Delta\tilde{p}_{t+m+1})\right] \end{aligned} \tag{5.98}$$

where $\tilde{y}^n_\infty \equiv \lim_{T\to\infty} E_t\left[\tilde{y}^n_{t+T}\right]$ is the expected long-run equilibrium output. In accordance with equation 5.90 this can be expressed in terms of the expected real interest rate

$$\tilde{c}^k_{t,h} = E_{t-k}\left[\tilde{y}^n_\infty - \sum_{m=0}^{\infty} \tilde{r}_{t+m}\right] \tag{5.99}$$

This can then be used with equation 5.97 in equation 5.73 to get

$$\begin{aligned}\tilde{w}^k_t &= E_{t-k}\left[\tilde{p}_t - \frac{\tilde{\gamma}_t}{\bar{\gamma}-1} + \tilde{c}_{t,h} + \frac{1}{\psi}(-\bar{\gamma}(\tilde{w}_{t,h} - \tilde{w}_t) + \tilde{l}_t\right] \\ &= E_{t-k}\left[\tilde{p}_t + \frac{\bar{\gamma}(\tilde{w}_t - \tilde{p}_t)}{\psi + \bar{\gamma}} - \frac{\tilde{\psi\gamma_t}}{(\bar{\gamma}-1)(\psi+\bar{\gamma})}\right. \\ &\qquad \left. + \frac{\psi(\tilde{y}^n_\infty - \sum_{m=0}^{\infty} \tilde{r}_{t+m})}{\psi + \bar{\gamma}} + \frac{\tilde{l}_t}{\psi + \bar{\gamma}}\right] \end{aligned} \tag{5.100}$$

Then for equation 5.76 use the linearized production function and demand/market-clearing condition (equations 5.77 and 5.92)

$$\begin{aligned}
\tilde{p}_t^j &= E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t + \tilde{n}_{t,f} - \tilde{y}_{t,f}\right] \\
&= E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t + \frac{\tilde{y}_{t,f} - \tilde{a}_t}{\beta} - \tilde{y}_{t,f}\right] \\
&= E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t + \frac{-\bar{\nu}(\tilde{p}_{t,f} - \tilde{p}_t) + \tilde{y}_t - \tilde{a}_t}{\beta} + \bar{\nu}(\tilde{p}_{t,f} - \tilde{p}_t) - \tilde{y}_t\right] \\
&= E_{t-j}\left[\frac{-\tilde{\nu}_t}{\bar{\nu}-1} + \tilde{w}_t - \frac{(1-\beta)\bar{\nu}(\tilde{p}_{t,f} - \tilde{p}_t)}{\beta} + \frac{(1-\beta)\tilde{y}_t}{\beta} - \frac{\tilde{a}_t}{\beta}\right] \\
\tilde{p}_t^j &= E_{t-j}\left[\frac{-\beta\tilde{\nu}_t/(\bar{\nu}-1) + \beta\tilde{w}_t + (1-\beta)\bar{\nu}\tilde{p}_t + (1-\beta)\tilde{y}_t - \tilde{a}_t}{\beta + \bar{\nu}(1-\beta)}\right]
\end{aligned} \tag{5.101}$$

Equations 5.99, 5.100 and 5.101 can now be used in equations 5.82, 5.85 and 5.88 as necessary to express aggregate price, consumption (or output) and wage as functions of aggregate variables, so that the model can be expressed entirely in aggregate terms. Using equation 5.99 in equation 5.85 gives

$$\tilde{c}_t = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_{t-k}\left[\tilde{y}_\infty^n - \sum_{m=0}^{\infty} \tilde{r}_{t+m}\right] \tag{5.102}$$

Using equation 5.100 in equation 5.88 gives

$$\begin{aligned}
\tilde{w}_t = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_{t-k}\Bigg[&\tilde{p}_t + \frac{\bar{\gamma}(\tilde{w}_t - \tilde{p}_t)}{\psi + \bar{\gamma}} - \frac{\tilde{\psi\gamma_t}}{(\bar{\gamma}-1)(\psi+\bar{\gamma})} \\
&+ \frac{\psi(\tilde{y}_\infty^n - \sum_{m=0}^{\infty} \tilde{r}_{t+m})}{\psi + \bar{\gamma}} + \frac{\tilde{l}_t}{\psi + \bar{\gamma}}\Bigg]
\end{aligned} \tag{5.103}$$

Using equation 5.101 in equation 5.82 gives

$$\tilde{p}_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \cdot$$
$$E_{t-j}\left[\frac{-\beta\tilde{\nu}_t/(\bar{\nu}-1) + \beta\tilde{w}_t + (1-\beta)\bar{\nu}\tilde{p}_t + (1-\beta)\tilde{y}_t - \tilde{a}_t}{\beta + \bar{\nu}(1-\beta)}\right] \tag{5.104}$$

Also, equation 5.102 can now be used in equation 5.79 to get

$$\tilde{y}_t = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_{t-k}\left[\tilde{y}_{\infty}^n - \sum_{m=0}^{\infty} \tilde{r}_{t+m}\right] + \tilde{g}_t \tag{5.105}$$

The approximate aggregate analogue to the production function (equation 5.77) is

$$\tilde{y}_t = \tilde{a}_t + \beta\tilde{l}_t \tag{5.106}$$

A.2.3 Model Solution

Now the five endogenous variables output, price, wage, labour and nominal interest rate are described by the five equations 5.91, 5.103, 5.104, 5.105 and 5.106. It remains to specify the exogenous processes, which are

$$\tilde{g}_t = \rho_g \tilde{g}_{t-1} + e_t^g \tag{5.107}$$
$$\tilde{\nu}_t = \rho_\nu \tilde{\nu}_{t-1} + e_t^\nu \tag{5.108}$$
$$\tilde{\gamma}_t = \rho_\gamma \tilde{\gamma}_{t-1} + e_t^\gamma \tag{5.109}$$
$$\tilde{\varepsilon}_t = \rho_\varepsilon \tilde{\varepsilon}_{t-1} + e_t^\varepsilon \tag{5.110}$$
$$\triangle\tilde{a}_t = \rho_{\Delta a} \triangle\tilde{a}_{t-1} + e_t^{\Delta a} \tag{5.111}$$

Furthermore $e_t^s \sim N(0, \sigma_s^2)$ and i.i.d. over time, $E\left[e_t^s e_{t+k}^s\right] = 0$ for $k \neq 0$ and $E\left[e_t^s e_t^{s'}\right] = 0$ for $s \neq s'$, $s \in (g, \nu, \gamma, \varepsilon, \triangle a)$.

The natural value of variables is defined here as the value which they would take if all agents had full information, i.e. $\lambda = \omega = 1$. Mankiw and Reis (2006b) show that the model equations in this case simplify and, in particular, give a natural output value as a function of the real exogenous variables as[16]

$$\tilde{y}_t^n = \tilde{a}_t + \frac{\psi\beta}{\psi+1}\left(\frac{\tilde{\gamma}_t}{\bar{\gamma}-1} + \tilde{g}_t + \frac{\tilde{\nu}_t}{\bar{\nu}-1}\right) \tag{5.112}$$

The solution method supposes a solution form of $z_t = \sum_{s\in S}\sum_{n=0}^{\infty}\hat{z}_n(s)e_{t-n}^s$ for $z_t \in Z_t = (\tilde{y}_t, \tilde{p}_t, \tilde{w}_t, \tilde{l}_t, \tilde{i}_t)$ and $s \in S = (g, \nu, \gamma, \varepsilon, \triangle a)$. $\hat{z}_n(s)$ is the response of variable z to a shock of type s occurring n periods ago. This supposed solution is then imposed on the model equations and the coefficients pertaining to each particular shock can be compared, which will lead to a system of linear difference equations which is amenable to numerical solution.

Define $\Lambda_n \equiv \lambda\sum_{k=0}^{n}(1-\lambda)^k$ and $\Omega_n \equiv \omega\sum_{k=0}^{n}(1-\omega)^k$.

Starting with shocks to monetary policy, substitution of the guessed solution into equation 5.104 gives

$$\sum_{n=0}^{\infty}\hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} = \lambda\sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\left[\frac{\beta\sum_{n=0}^{\infty}\hat{w}_n(\varepsilon)e_{t-n}^{\varepsilon}}{\beta+\bar{\nu}(1-\beta)}\right.$$
$$\left.+\frac{(1-\beta)\bar{\nu}\sum_{n=0}^{\infty}\hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} + (1-\beta)\sum_{n=0}^{\infty}\hat{y}_n(\varepsilon)e_{t-n}^{\varepsilon}}{\beta+\bar{\nu}(1-\beta)}\right] \tag{5.113}$$

To allow for the various expectations, note that

$$\lambda\sum_{j=0}^{\infty}(1-\lambda)^j\sum_{n=j}^{\infty}\hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} = \sum_{n=0}^{\infty}\lambda\sum_{j=0}^{n}(1-\lambda)^j\hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} \tag{5.114}$$

[16]Mankiw and Reis (2006b)p17.

Hence equation 5.113 can be rewritten as

$$\Big\{\beta + \bar{\nu}(1-\beta)\Big\}\sum_{n=0}^{\infty}\hat{p}_n(\varepsilon)e^{\varepsilon}_{t-n} =$$
$$\sum_{n=0}^{\infty}\Lambda_n\Big\{\beta\hat{w}_n(\varepsilon) + \bar{\nu}(1-\beta)\hat{p}_n(\varepsilon) + (1-\beta)\hat{y}_n(\varepsilon)\Big\}e^{\varepsilon}_{t-n} \tag{5.115}$$

Then for any particular period $t-n$ the coefficients must match, so

$$\Big\{\frac{\beta + \bar{\nu}(1-\beta)}{\Lambda_n} - \bar{\nu}(1-\beta)\Big\}\hat{p}_n(\varepsilon) = \beta\hat{w}_n(\varepsilon) + (1-\beta)\hat{y}_n(\varepsilon) \tag{5.116}$$

The intention is to derive a difference equation in the price responses, $\hat{p}_n(\varepsilon)$, so it is necessary to use the other equations to replace $\hat{w}_n(\varepsilon)$ and $\hat{y}_n(\varepsilon)$. From equation 5.105

$$\sum_{n=0}^{\infty}\hat{y}_n(\varepsilon)e^{\varepsilon}_{t-n} = \omega\sum_{k=0}^{\infty}(1-\omega)^k E_{t-k}\Big[-\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\hat{r}_n(\varepsilon)e^{\varepsilon}_{t+m-n}\Big] \tag{5.117}$$

Then by the same reasoning as for equation 5.114 this becomes

$$\sum_{n=0}^{\infty}\hat{y}_n(\varepsilon)e^{\varepsilon}_{t-n} = \omega\sum_{k=0}^{\infty}(1-\omega)^k\Big(-\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon)e^{\varepsilon}_{t-n}\Big)$$
$$\sum_{n=0}^{\infty}\hat{y}_n(\varepsilon)e^{\varepsilon}_{t-n} = \sum_{n=0}^{\infty}\Omega_n\Big(-\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon)e^{\varepsilon}_{t-n}\Big) \tag{5.118}$$

and for any period $t-n$ this implies

$$\hat{y}_n(\varepsilon) = -\Omega_n\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon) \tag{5.119}$$

Then from equation 5.103

$$\sum_{n=0}^{\infty}\hat{w}_n(\varepsilon)e^{\varepsilon}_{t-n} = \omega\sum_{k=0}^{\infty}(1-\omega)^k\Bigg\{\sum_{n=k}^{\infty}\hat{p}_n(\varepsilon)e^{\varepsilon}_{t-n} + \frac{\bar{\gamma}}{\psi+\bar{\gamma}}\Bigg(\sum_{n=k}^{\infty}\hat{w}_n(\varepsilon)e^{\varepsilon}_{t-n} - \sum_{n=k}^{\infty}\hat{p}_n(\varepsilon)e^{\varepsilon}_{t-n}\Bigg) + \frac{1}{\beta(\psi+\bar{\gamma})}\sum_{n=k}^{\infty}\hat{y}_n(\varepsilon)e^{\varepsilon}_{t-n} + \frac{\psi}{\psi+\bar{\gamma}}\Bigg(-\sum_{n=k}^{\infty}\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon)e^{\varepsilon}_{t-n}\Bigg)\Bigg\} \tag{5.120}$$

which can then be re-arranged to

$$\sum_{n=0}^{\infty}\hat{w}_n(\varepsilon)e^{\varepsilon}_{t-n} = \sum_{n=0}^{\infty}\omega\sum_{k=0}^{\infty}(1-\omega)^k\Bigg\{\hat{p}_n(\varepsilon) + \frac{\bar{\gamma}}{\psi+\bar{\gamma}}\Big(\hat{w}_n(\varepsilon)-\hat{p}_n(\varepsilon)\Big) + \frac{\hat{y}_n(\varepsilon)}{\beta(\psi+\bar{\gamma})} - \frac{\psi}{\psi+\bar{\gamma}}\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon)\Bigg\}e^{\varepsilon}_{t-n} \tag{5.121}$$

and therefore

$$\hat{w}_n(\varepsilon) = \frac{1}{\psi+\bar{\gamma}(1-\Omega_n)}\Bigg(\psi\Omega_n\hat{p}_n(\varepsilon) + \Omega_n\frac{\hat{y}_n(\varepsilon)}{\beta} - \psi\Omega_n\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon)\Bigg) \tag{5.122}$$

From equation 5.119 $-\sum_{m=0}^{\infty}\hat{r}_{n+m}(\varepsilon) = \frac{\hat{y}_n(\varepsilon)}{\Omega_n}$. Using this in equation 5.122 and the result in equation 5.116 gives

$$\Big\{\frac{\beta+\bar{\nu}(1-\beta)}{\Lambda_n} - \bar{\nu}(1-\beta)\Big\}\hat{p}_n(\varepsilon) = \beta\Bigg\{\frac{1}{\psi+\bar{\gamma}(1-\Omega_n)}\Bigg(\psi\Omega_n\hat{p}_n(\varepsilon) + \Big(\frac{\Omega_n}{\beta}+\psi\Big)\hat{y}_n(\varepsilon)\Bigg)\Bigg\} + (1-\beta)\hat{y}_n(\varepsilon) \tag{5.123}$$

Re-arrangement of this equation results in

$$\hat{y}_n(\varepsilon) = \Psi_n \hat{p}_n(\varepsilon) \tag{5.124}$$

where[17]

$$\begin{aligned}
\Psi_n &= \frac{\Psi_n^{num}}{\Psi_n^{den}} \\
&= \frac{\Omega_n\left[\left\{\psi + \bar{\gamma}(1-\Omega_n)\right\}\left\{\frac{\beta+\bar{\nu}(1-\beta)}{\Lambda_n} - \bar{\nu}(1-\beta)\right\} - \beta\psi\Omega_n\right]}{\left[(1-\beta)\Omega_n(\psi+\bar{\gamma}) + \Omega_n\left\{\Omega_n(1-\bar{\gamma}(1-\beta)) + \beta\psi\right\}\right]}
\end{aligned} \tag{5.125}$$

The next step is to use equations 5.90 and 5.91[18]

$$\begin{aligned}
\widetilde{i}_t = \widetilde{r}_t + E_t\left[\tilde{p}_{t+1} - \tilde{p}_t\right] &= \phi_y(\tilde{y}_t - \tilde{y}_t^n) + \phi_p(\tilde{p}_t - \tilde{p}_{t-1}) - \varepsilon_t \\
\sum_{n=0}^{\infty} \hat{r}_n(\varepsilon)e_{t-n}^{\varepsilon} + \sum_{n=0}^{\infty} \hat{p}_{n+1}(\varepsilon)e_{t-n}^{\varepsilon} - \sum_{n=0}^{\infty} \hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} &= \\
\phi_y \sum_{n=0}^{\infty} \hat{y}_n(\varepsilon)e_{t-n}^{\varepsilon} + \phi_p\left(\sum_{n=0}^{\infty} \hat{p}_n(\varepsilon)e_{t-n}^{\varepsilon} - \sum_{n=0}^{\infty} \hat{p}_{n-1}(\varepsilon)e_{t-n}^{\varepsilon}\right)& \\
-\sum_{n=0}^{\infty} \rho_{\varepsilon}^n e_{t-n}^{\varepsilon}&
\end{aligned} \tag{5.126}$$

For any individual n

$$\begin{aligned}
\hat{r}_n(\varepsilon) + \hat{p}_{n+1}(\varepsilon) - \hat{p}_n(\varepsilon) &= \phi_y\hat{y}_n(\varepsilon) + \phi_p(\hat{p}_n(\varepsilon) - \hat{p}_{n-1}(\varepsilon)) - \rho_{\varepsilon}^n \\
\phi_p\hat{p}_{n-1}(\varepsilon) &= \phi_y\hat{y}_n(\varepsilon) + (1+\phi_p)\hat{p}_n(\varepsilon) - \hat{p}_{n+1}(\varepsilon) - \hat{r}_n(\varepsilon) - \rho_{\varepsilon}^n
\end{aligned} \tag{5.127}$$

Now it is desired to replace the output and real interest rate responses with expressions involving the price response coefficients. Beginning with the real interest rate, from

[17]The re-arrangement is chosen to match the expression in MR p18.

[18]Recall that the natural level of output does not depend upon monetary policy shocks.

equation 5.119

$$\begin{aligned}\frac{\hat{y}_n(\varepsilon)}{\Omega_n} &= -\sum_{m=0}^{\infty} \hat{r}_{n+m}(\varepsilon) \\ \therefore \frac{\hat{y}_{n+1}(\varepsilon)}{\Omega_{n+1}} &= -\sum_{m=0}^{\infty} \hat{r}_{n+1+m}(\varepsilon) \\ \Rightarrow \frac{\hat{y}_{n+1}(\varepsilon)}{\Omega_{n+1}} - \frac{\hat{y}_n(\varepsilon)}{\Omega_n} &= \hat{r}_n(\varepsilon) \qquad (5.128)\end{aligned}$$

so equation 5.127 becomes

$$\begin{aligned}\phi_p \hat{p}_{n-1}(\varepsilon) = \phi_y \hat{y}_n(\varepsilon) + (1+\phi_p)\hat{p}_n(\varepsilon) - \hat{p}_{n+1}(\varepsilon) \\ -\frac{\hat{y}_{n+1}(\varepsilon)}{\Omega_{n+1}} + \frac{\hat{y}_n(\varepsilon)}{\Omega_n} - \rho_\varepsilon^n \qquad (5.129)\end{aligned}$$

and now equation 5.124 can be used to achieve an equation involving only the price response coefficients (and a constant)

$$\begin{aligned}\phi_p \hat{p}_{n-1}(\varepsilon) = \phi_y \Psi_n \hat{p}_n(\varepsilon) + (1+\phi_p)\hat{p}_n(\varepsilon) - \hat{p}_{n+1}(\varepsilon) \\ -\frac{\Psi_{n+1}\hat{p}_{n+1}(\varepsilon)}{\Omega_{n+1}} + \frac{\Psi_n \hat{p}_n(\varepsilon)}{\Omega_n} - \rho_\varepsilon^n \qquad (5.130)\end{aligned}$$

and this re-arranges to

$$\begin{aligned}\left(1 + \frac{\Psi_{n+1}}{\Omega_{n+1}}\right)\hat{p}_{n+1}(\varepsilon) - \left(1 + \phi_p + \frac{\Psi_n}{\Omega_n} + \phi_y \Psi_n\right)\hat{p}_n(\varepsilon) + \phi_p \hat{p}_{n-1}(\varepsilon) \\ = -\rho_\varepsilon^n \qquad (5.131)\end{aligned}$$

With the initial condition $\hat{p}_{-1}(\varepsilon) = 0$ and terminal condition $\hat{p}_N(\varepsilon) = 0$[19] the price response coefficients can be found as the solution to the system of equations[20]

[19]The same conditions will be used for all $s \in S$, with s instead of ε. The terminal condition is an approximation which becomes increasingly appropriate as N increases.

[20]Mankiw and Reis (2006b) p20.

$$\begin{pmatrix} -B_0 & A_1 & \dots & 0 & 0 & 0 \\ \phi_p & -B_1 & A_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \phi_p & -B_{N-2} & A_{N-1} & 0 \\ 0 & 0 & 0 & \phi_p & -B_{N-1} & A_N \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \hat{p}_0(\varepsilon) \\ \hat{p}_1(\varepsilon) \\ \dots \\ \hat{p}_{N-2}(\varepsilon) \\ \hat{p}_{N-1}(\varepsilon) \\ \hat{p}_N(\varepsilon) \end{pmatrix} = \begin{pmatrix} -\rho_\varepsilon^0 \\ -\rho_\varepsilon^1 \\ \dots \\ -\rho_\varepsilon^{N-2} \\ -\rho_\varepsilon^{N-1} \\ -\rho_\varepsilon^N \end{pmatrix} \tag{5.132}$$

where $A_n = 1 + \frac{\Psi_n}{\Omega_n}$ and $B_n = A_n + \phi_y \Psi_n + \phi_p$.

Having obtained a solution for the responses of price to monetary policy shocks it is then necessary to derive the responses of price to each of the four other shocks. Rather than repeat the above twenty-five equations four times, the following states only those significant equations where there is a difference (other than $s \neq \varepsilon$) to the above, with the implication being that the use of the equations below in the appropriate steps of the solution sequence above will give the relevant solution. Unless otherwise stated all definitions from above apply equally below.

For aggregate demand/fiscal shocks, $s = g$, equation 5.119 is replaced by

$$\hat{y}_n(g) = -\Omega_n \sum_{m=0}^{\infty} \hat{r}_{n+m}(g) + \rho_g^n \tag{5.133}$$

and equation 5.124 is now

$$\hat{y}_n(g) = \Psi_n \hat{p}_n(g) + \frac{\beta\psi\Omega_n}{\Psi_n^{den}} \rho_g^n \tag{5.134}$$

All shocks other than monetary policy shocks affect the natural level of output, so there is a different term in the equivalent of equation 5.127

$$\phi_p \hat{p}_{n-1}(g) = \phi_y \hat{y}_n(g) + (1 + \phi_p)\hat{p}_n(g) - \hat{p}_{n+1}(g) - \hat{r}_n(g) - \frac{\phi_y \psi \beta}{\psi + 1} \rho_g^n \tag{5.135}$$

The real interest rate responses can be substituted using

$$\frac{\hat{y}_{n+1}(g)-\rho_g^{n+1}}{\Omega_{n+1}}-\frac{\hat{y}_n(g)-\rho_g^n}{\Omega_n}=\hat{r}_n(g) \tag{5.136}$$

and replacing the output responses using equation 5.134 gives the difference equation

$$\begin{aligned}\left(1+\frac{\Psi_{n+1}}{\Omega_{n+1}}\right)\hat{p}_{n+1}(g)-\left(1+\phi_p+\frac{\Psi_n}{\Omega_n}+\phi_y\Psi_n\right)\hat{p}_n(g)+\phi_p\hat{p}_{n-1}(g)= \\ \left[\phi_y\left(\Upsilon_n(g)-\frac{\beta\psi}{\psi+1}\right)+\frac{1-\Upsilon_{n+1}(g)}{\Omega_{n+1}}\rho_g-\frac{1-\Upsilon_n(g)}{\Omega_n}\right]\rho_g^n\end{aligned} \tag{5.137}$$

where $\Upsilon_n(g)\equiv\frac{\beta\psi\Omega_n}{\Psi_n^{den}}$. The solutions can then be found using a system equivalent to equation 5.132.

Next, consider $s=\nu$. Equation 5.124 becomes

$$\hat{y}_n(\nu)=\Psi_n\hat{p}_n(\nu)+\frac{\beta\Big\{\psi+\bar{\gamma}(1-\Omega_n)\Big\}\Omega_n}{(\bar{\nu}-1)\Psi_n^{den}}\rho_\nu^n \tag{5.138}$$

Equation 5.129 is now

$$\begin{aligned}\phi_p\hat{p}_{n-1}(\nu)=\phi_y\hat{y}_n(\nu)+(1+\phi_p)\hat{p}_n(\nu)-\hat{p}_{n+1}(\nu) \\ -\frac{\hat{y}_{n+1}(\nu)}{\Omega_{n+1}}+\frac{\hat{y}_n(\nu)}{\Omega_n}-\frac{\phi_y\psi\beta}{(\psi+1)(\bar{\nu}-1)}\rho_\nu^n\end{aligned} \tag{5.139}$$

and this leads to the difference equation

$$\begin{aligned}\left(1+\frac{\Psi_{n+1}}{\Omega_{n+1}}\right)\hat{p}_{n+1}(\nu)-\left(1+\phi_p+\frac{\Psi_n}{\Omega_n}+\phi_y\Psi_n\right)\hat{p}_n(\nu)+\phi_p\hat{p}_{n-1}(\nu)= \\ \left[\phi_y\left(\Upsilon_n(\nu)-\frac{\beta\psi}{(\psi+1)(\bar{\nu}-1)}\right)-\frac{\Upsilon_{n+1}(\nu)}{\Omega_{n+1}}\rho_\nu+\frac{\Upsilon_n(\nu)}{\Omega_n}\right]\rho_\nu^n\end{aligned} \tag{5.140}$$

with $\Upsilon_n(\nu)\equiv\frac{\beta\Big\{\psi+\bar{\gamma}(1-\Omega_n)\Big\}\Omega_n}{(\bar{\nu}-1)\Psi_n^{den}}$. Again, the solutions for the price response coefficients can then be found by solving an equivalent to equation 5.132.

The next shock type is e_t^{γ}. Here the only differences to the case of e_t^{ε} is that there are extra terms in the wage index equation and in the equation for the natural level of output. As a result equation 5.124 becomes

$$\hat{y}_n(\gamma) = \Psi_n \hat{p}_n(\gamma) + \frac{\beta\psi\Omega_n^2}{(\bar{\gamma}-1)\Psi_n^{den}}\rho_\gamma^n \tag{5.141}$$

and equation 5.131 is instead

$$\left(1+\frac{\Psi_{n+1}}{\Omega_{n+1}}\right)\hat{p}_{n+1}(\gamma) - \left(1+\phi_p+\frac{\Psi_n}{\Omega_n}+\phi_y\Psi_n\right)\hat{p}_n(\gamma) + \phi_p\hat{p}_{n-1}(\gamma) =$$
$$\left[\phi_y\left(\Upsilon_n(\gamma) - \frac{\beta\psi}{(\psi+1)(\bar{\gamma}-1)}\right) - \frac{\Upsilon_{n+1}(\gamma)}{\Omega_{n+1}}\rho_\gamma + \frac{\Upsilon_n(\gamma)}{\Omega_n}\right]\rho_\gamma^n \tag{5.142}$$

with $\Upsilon_n(\gamma) \equiv \frac{\beta\psi\Omega_n^2}{(\bar{\gamma}-1)\Psi_n^{den}}$.

The last shock type is the technology shock. The technology process is $\tilde{a}_t = \tilde{a}_{t-1} + \Delta\tilde{a}_t$, so by repeated substitution and with an assumed initial value of zero, $\tilde{a}_t = \sum_{j=0}^{\infty}\Delta\tilde{a}_{t-j}$. Given the specification of the technology shocks in equation 5.111, this can be written as $\tilde{a}_t = \sum_{j=0}^{\infty}\sum_{n=0}^{\infty}\rho_{\Delta a}^n e_{t-n-j}^{\Delta a} = \sum_{n=0}^{\infty}\sum_{j=0}^{n}\rho_{\Delta a}^j e_{t-n}^{\Delta a}$. The solution then proceeds as for the other shocks. Equation 5.116 becomes

$$\left\{\frac{\beta+\bar{\nu}(1-\beta)}{\Lambda_n} - \bar{\nu}(1-\beta)\right\}\hat{p}_n(\Delta a) = \beta\hat{w}_n(\Delta a) + (1-\beta)\hat{y}_n(\Delta a) - \sum_{j=0}^{n}\rho_{\Delta a}^j \tag{5.143}$$

and equation 5.122 becomes

$$\hat{w}_n(\Delta a) = \frac{1}{\psi+\bar{\gamma}(1-\Omega_n)}\Bigg(\psi\Omega_n\hat{p}_n(\Delta a) +$$
$$\Omega_n\left(\frac{1}{\beta}+\frac{\psi}{\Omega_n}\right)\hat{y}_n(\Delta a) - \frac{\Omega_n}{\beta}\sum_{j=0}^{n}\rho_{\Delta a}^j\Bigg) \tag{5.144}$$

These equations lead to

$$\hat{y}_n(\Delta a) = \Psi_n \hat{p}_n(\Delta a) + \frac{\Omega_n \left\{ \psi + \bar{\gamma} + \Omega_n(1-\bar{\gamma}) \right\}}{\Psi_n^{den}} \sum_{j=0}^{n} \rho_{\Delta a}^{j} \tag{5.145}$$

Noting that $\sum_{j=0}^{n} \rho_{\Delta a}^{j} = \frac{1-\rho_{\Delta a}^{n+1}}{1-\rho_{\Delta a}}$, the use of the above equation with the monetary policy rule gives

$$\begin{aligned} \left(1 + \frac{\Psi_{n+1}}{\Omega_{n+1}}\right)\hat{p}_{n+1}(\Delta a) - \left(1 + \phi_p + \frac{\Psi_n}{\Omega_n} + \phi_y \Psi_n\right)\hat{p}_n(\Delta a) \\ + \phi_p \hat{p}_{n-1}(\Delta a) = \\ \frac{1}{1-\rho_{\Delta a}}\left[(1-\rho_{\Delta a}^{n+1})\left\{\phi_y\Big(\Upsilon_n(\Delta a) - 1\Big) + \frac{\Upsilon_n(\Delta a)}{\Omega_n}\right\}\right. \\ \left. -(1-\rho_{\Delta a}^{n+2})\frac{\Upsilon_{n+1}(\Delta a)}{\Omega_{n+1}}\right] \end{aligned} \tag{5.146}$$

with $\Upsilon_n(\Delta a) \equiv \frac{\Omega_n \left\{ \psi + \bar{\gamma} + \Omega_n(1-\bar{\gamma}) \right\}}{\Psi_n^{den}}$.

Overall, the solution for price as a sum of responses to the shocks can be found from the five equations 5.131, 5.137, 5.140, 5.142 and 5.146, by solving the system in equation 5.132 (or its equivalent).

Once this solution is found, the intermediate equations given in this section can be used to find the solutions for the other endogenous variables:

$$
\begin{aligned}
\hat{y}_n(\varepsilon) &= \Psi_n \hat{p}_n(\varepsilon) \\
\hat{y}_n(g) &= \Psi_n \hat{p}_n(g) + \Upsilon_n(g)\rho_g^n \\
\hat{y}_n(\nu) &= \Psi_n \hat{p}_n(\nu) + \Upsilon_n(\nu)\rho_\nu^n \\
\hat{y}_n(\gamma) &= \Psi_n \hat{p}_n(\gamma) + \Upsilon_n(\gamma)\rho_\gamma^n \\
\hat{y}_n(\Delta a) &= \Psi_n \hat{p}_n(\Delta a) + \Upsilon_n(\Delta a)\frac{1-\rho_{\Delta a}^{n+1}}{1-\rho_{\Delta a}}
\end{aligned}
$$

$$
\begin{aligned}
\hat{r}_n &= \frac{\hat{y}_{n+1}(s)}{\Omega_{n+1}} - \frac{\hat{y}_n(s)}{\Omega_n} \quad s \in (\varepsilon, \nu, \gamma, \Delta a) \\
\hat{r}_n &= \frac{\hat{y}_{n+1}(g) - \rho_g^{n+1}}{\Omega_{n+1}} - \frac{\hat{y}_n(g) - \rho_g^n}{\Omega_n} \\
\hat{i}_n(s) &= \hat{r}_n(s) + \hat{p}_{n+1}(s) - \hat{p}_n(s) \quad s \in S
\end{aligned}
$$

$$
\begin{aligned}
\hat{l}_n(s) &= \frac{\hat{y}_n(s)}{\beta} \quad s \in (\varepsilon, g, \nu, \gamma) \\
\hat{l}_n(\Delta a) &= \frac{\hat{y}_n(\Delta a)}{\beta} - \frac{1-\rho_{\Delta a}^{n+1}}{\beta(1-\rho_{\Delta a})} \\
\hat{w}_n(s) &= \frac{1}{\beta}\left[\left\{\frac{\beta + \bar{\nu}(1-\beta)}{\Omega_n} - \bar{\nu}(1-\beta)\right\}\hat{p}_n(s) - (1-\beta)\hat{y}_n(s)\right] \quad s \in (\varepsilon, g, \gamma) \\
\hat{w}_n(\nu) &= \frac{1}{\beta}\left[\left\{\frac{\beta + \bar{\nu}(1-\beta)}{\Omega_n} - \bar{\nu}(1-\beta)\right\}\hat{p}_n(\nu) - (1-\beta)\hat{y}_n(\nu) + \frac{\beta\rho_\nu^n}{\bar{\nu}-1}\right] \\
\hat{w}_n(\Delta a) &= \frac{1}{\beta}\left[\left\{\frac{\beta + \bar{\nu}(1-\beta)}{\Omega_n} - \bar{\nu}(1-\beta)\right\}\hat{p}_n(\Delta a) - (1-\beta)\hat{y}_n(\Delta a) + \frac{1-\rho_{\Delta a}^{n+1}}{1-\rho_{\Delta a}}\right]
\end{aligned}
$$

Bibliography

Ball, L., Mankiw, N., Reis, R., 2005. Monetary policy for inattentive economies. JOURNAL OF MONETARY ECONOMICS 52, 703–725. Research Conference on Behavioral Macroeconomics, Gerzensee, SWITZERLAND, OCT 11-12, 2002.

Ball, L., Romer, D., 1990. Real Rigidities And The Nonneutrality Of Money. REVIEW OF ECONOMIC STUDIES 57, 183–203.

Branch, W.A., Carlson, J., Evans, G.W., McGough, B., 2009. Monetary Policy, Endogenous Inattention And The Volatility Trade-Off. ECONOMIC JOURNAL 119, 123–157.

Branch, W.A., McGough, B., 2009. A New Keynesian model with heterogeneous expectations. JOURNAL OF ECONOMIC DYNAMICS & CONTROL 33, 1036–1051. Workshop on Complexity in Economics and Finance, Leiden, NETHERLANDS, OCT 22-27, 2007.

Calvo, G., 1983. Staggered Prices In A Utility-Maximizing Framework. JOURNAL OF MONETARY ECONOMICS 12, 383–398.

Canova, F., 2007. Methods for Applied Macroeconomic Research. Princeton University Press.

Canova, F., De Nicolo, G., 2002. Monetary disturbances matter for business fluctuations in the G-7. JOURNAL OF MONETARY ECONOMICS 49, 1131–1159.

Carrillo, J.A., 2012. How well does sticky information explain the dynamics of inflation, output, and real wages? JOURNAL OF ECONOMIC DYNAMICS & CONTROL 36, 830–850.

Clarida, Gali, J., Gertler, M., 2000. Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. Quarterly Journal of Economics 115, 147–180.

Coibion, O., 2010. Testing The Sticky Information Phillips Curve. REVIEW OF ECONOMICS AND STATISTICS 92, 87–101.

Coibion, O., Gorodnichenko, Y., 2011. Strategic Interaction Among Heterogeneous Price-Setters In An Estimated DSGE Model. Review of Economics and Statistics 93, 920–940.

Cone, T.E., 2008. Optimal information acquisition and monetary policy. JOURNAL OF MACROECONOMICS 30, 1370–1389.

Cooper, R., John, A., 1988. Coordinating Coordination Failures In Keynesian Models. QUARTERLY JOURNAL OF ECONOMICS 103, 441–463.

Cooper, R.W., 1999. Coordination Games: Complementarities and Macroeconomics. Cambridge University Press.

Drager, L., 2011. Endogenous Persistence with Recursive Inattentiveness. DEP Discussion Papers - Macroeconomics and Finance Series 3/2011 .

Dupor, B., Kitamura, T., Tsuruga, T., 2010. Integrating Sticky Prices And Sticky Information. REVIEW OF ECONOMICS AND STATISTICS 92, 657–669.

Erceg, C., Henderson, D., Levin, A., 2000. Optimal monetary policy with staggered wage and price contracts. JOURNAL OF MONETARY ECONOMICS 46, 281–313.

Evans, G., Honkapohja, S., Marimon, R., 2001. Convergence in monetary inflation models with heterogeneous learning rules. MACROECONOMIC DYNAMICS 5, 1–31.

Evans, G.W., Honkapohja, S., 2009. Learning and Macroeconomics. ANNUAL REVIEW OF ECONOMICS 1, 421–449.

Fry, R., Pagan, A., 2011. Sign Restrictions in Structural Vector Autoregressions: A Critical Review. JOURNAL OF ECONOMIC LITERATURE 49, 938–960.

Fukac, M., Pagan, A., 2010. LIMITED INFORMATION ESTIMATION AND EVALUATION OF DSGE MODELS. JOURNAL OF APPLIED ECONOMETRICS 25, 55–70. Workshop on Model Uncertainty and Macroeconomics, Wellington, NEW ZEALAND, JUN 27-28, 2006.

Fukunaga, I., 2007. Imperfect common knowledge, staggered price setting, and the effects of monetary policy. JOURNAL OF MONEY CREDIT AND BANKING 39, 1711–1739.

Gali, J., 2008. Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton University Press.

Garratt, A., Robertson, D., Wright, S., 2006. Permanent vs transitory components and economic fundamentals. JOURNAL OF APPLIED ECONOMETRICS 21, 521–542.

Heer, B., Maussner, A., 2009. Dynamic General Equilibrium Modeling: Computational Methods and Applications. Springer.

Hellwig, C., Veldkamp, L., 2009. Knowing What Others Know: Coordination Motives in Information Acquisition. REVIEW OF ECONOMIC STUDIES 76, 223–251.

Knotek, E.S., 2010. A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment. JOURNAL OF MONEY CREDIT AND BANKING 42, 1543–1564.

Lutkepohl, H., 2006. New Introduction to Multiple Time Series Analysis. Springer.

Mackowiak, B., Wiederholt, M., 2009. Optimal Sticky Prices under Rational Inattention. AMERICAN ECONOMIC REVIEW 99, 769–803.

Mankiw, N., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. QUARTERLY JOURNAL OF ECONOMICS 117, 1295–1328.

Mankiw, N., Reis, R., 2006b. Sticky Information in General Equilibrium. NBER Working Paper 12605 .

Mankiw, N.G., Reis, R., 2006a. Pervasive stickiness. AMERICAN ECONOMIC REVIEW 96, 164–169. 118th Annual Meeting of the American-Economic-Association, Boston, MA, JAN 06-08, 2006.

Mankiw, N.G., Reis, R., 2007. Sticky information in general equilibrium. JOURNAL OF THE EUROPEAN ECONOMIC ASSOCIATION 5, 603–613. 21th Annual Congress of the European-Economic-Association, Vienna, AUSTRIA, AUG 24-28, 2006.

McCandless, G., 2008. The ABCs of RBCs: An Introduction to Dynamic Macroeconomic Models. Harvard University Press.

Meyer-Gohde, A., 2010. Linear rational-expectations models with lagged expectations: A synthetic method. JOURNAL OF ECONOMIC DYNAMICS & CONTROL 34, 984–1002.

Paustian, M., Pytlarczyk, E., 2006. Sticky Contracts or Sticky Information? Evidence From An Estimated Euro Area DSGE Model. Bank of Canada Conference 'Dynamic Models Useful For Policy Making' Conference Paper.

Reis, R., 2009. Optimal Monetary Policy Rules in an Estimated Sticky-Information Model. American Economic Journal: Macroeconomics 1, 1–28.

Rudebusch, G., 2002. Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia. JOURNAL OF MONETARY ECONOMICS 49, 1161–1187.

Ruge-Murcia, F.J., 2007. Methods to estimate dynamic stochastic general equilibrium models. JOURNAL OF ECONOMIC DYNAMICS & CONTROL 31, 2599–2636.