Learning in Macroeconomics: An Empirical Approach

Inaugural–Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts– und Gesellschaftswissenschaften
durch die
Rechts– und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von
Simon Junker
aus Bielefeld

Bonn 2009

Dekan:	Prof. Dr. Christian Hillgruber
Erstreferent:	Prof. Dr. Jürgen von Hagen
Zweitreferent:	Prof. Dr. Ludger Linnemann

Tag der mündlichen Prüfung: 6. April 2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Acknowledgements

This thesis presents the main results of my research at the Bonn Graduate School of Economics (BGSE) and the Center for European Integration Studies (ZEI). I am indebted to a number of people who supported me and my work during the past years. Firstly, I would like to thank my supervisor Prof. Dr. Jürgen von Hagen for his guidance and support. I benefited much from his comments and the stimulating research environment he provided at the ZEI. Especially important for my studies were the yearly ZEI summer schools on Monetary Theory and Policy. I also want to thank my colleagues Oliver Fries, Michael Evers and Michael Massmann for numerous fruitful discussions on various economic and econometric topics. I am grateful to all colleagues from the University of Bonn, the participants of the BGSE workshops and the ZEI summer schools for various comments and suggestions. Moreover, I want to thank Prof. Dr. Urs Schweizer, Prof. Dr. Georg Nöldecke and again my supervisor Prof. Dr. Jürgen von Hagen for their combined efforts in running the BGSE. The support of the BGSE, the DFG and the ZEI is also gratefully acknowledged.

Contents

Introduction

Expectations play a central role in dynamic macroeconomic models. The standard modelling assumption, originating in the work of Muth (1961) and tracing back to Lucas' (1972) and (1976) seminal contributions, is the rational expectations hypothesis, which is frequently referred to as a revolution in macroeconomic thinking for the essential impact it has exerted on economic research, cf. e.g. Taylor (1999). The key idea is that agents' expectations are consistent with the forecast model derived from the underlying economic structure, implying that they make rational use of their knowledge about economic relationships and only incur non–systematic, unavoidable forecast errors. Hence, in the same way that microfounded macro models presuppose maximizing behavior in agents' allocations, the rational expectations hypothesis insists on optimal behavior in information processing.

The strong informational requirements of the rational expectations approach, however, have been subject to criticism, since implementing such forecasts requires the knowledge not only of the underlying model but also of the exact values of all of its parameters. A recent approach that moderates these strong prerequisites is the theory of adaptive learning. In its basic form it retains the assumption of agents knowing the equations of the economic model, but requires them to behave like econometricians to infer estimates of the model's parameters.

Adaptive learning thus keeps track of the evolution of agents' beliefs, i.e. the set of parameter estimates they base their forecasts on. Given such a set of estimates, agents can use the model's forecast equations to project the future evolution of the relevant variables. The literature has coined the term "perceived law of motion" to distinguish the agents' forecast model from the actual outcomes, which are referred to as the "actual law of motion" implied by both the economic structure and the process of forming expectations.

In this context, the concept of rational expectations calls for the consistency of perceptions and actual outcomes. Expectations formulae based on the agents' forecast model imply actual economic outcomes. These dynamic equations shall correspond to the systems perceived by agents, from which they derived their forecast model. Technically, the economic structure induces a mapping from perceived onto actual laws of motion, and the rational expectations equilibrium can be defined as a fixed point of this relation.

Much research has been devoted to solve this fixed point problem and the related problem of uniqueness. Important contributions include Blanchard and Kahn (1980), Klein (2000) and much of the research of McCallum. This line of research developed solution methods and elaborated conditions that guarantee the uniqueness of the rational expectations equilibrium. These conditions can be translated into policy prescriptions, for example the famous Taylor (1993) principle that states that only a sufficiently strong interest rate reaction to inflation on part of a central bank yields a unique equilibrium, precluding the existence of sunspots and self–fulfilling prophecies.

The adaptive learning approach conjectures that agents continuously observe their economic environment and use standard econometric procedures to fine–tune their understanding of dynamic relations. At any given point in time agents make optimal allocation decisions which crucially depend on their expectations about future circumstances. To determine the potential evolution of key variables they dispose only of an approximate model that is subject to statistical errors. The structural equations links their current perceived law of motion onto an actual law of motion, and the latter will change unless the economy has reached the rational expectations equilibrium. During the transition period, agents accumulate information that is contaminated by their imperfect knowledge, hence creating a feedback.

A natural question that arises in this context is whether the perceptions which in general are initially wrong will eventually move to the correct underlying full information equilibrium. If information is initially imperfect, will the economy nevertheless move towards the full information equilibrium or is the economic structure such that initial deviations from this state will accumulate and drive the system further apart from it? The work of Evans and Honkapohja (2001) provides the necessary theoretical background to analyze this convergence issue, known as E–stability, and opens a fruitful strand of research. Correspondingly, much work has been devoted to derive conditions on policy that assure E–stability,

Adaptive learning can isolate a unique equilibrium in a multiple equilibrium setup that is characterized by stability properties in the process of self–referential data accumulation, and is therefore often seen as an equilibrium selection device. The rational expectations theory cannot explain the transition from a prevailing equilibrium to a new one once structural factors have changed; instead it implies an immediate adaption to the new state without explaining how and whether the economy will get there. Adaptive learning provides this missing link in explicitly accounting for this transition period.

This new approach is, however, also capable of explaining many empirical phenomena. If the economic structure remains constant, rational expectations theory implies that the actual law of motion remains constant as well, while the information accumulation process, reflected in ever updating perceived laws of motion, will necessarily introduce time variation into actual economic dynamics. Hence, even if a theoretical model covers sufficient phenomena to imply a time invariant structure, we will nevertheless observe time variation over the course of learning the precise working of the model. Consequently, changes in econometric estimates (e.g. even changes in policy, see chapter 1) that are usually interpreted as exogenous changes in the structure that cannot be explained by the model can potentially be endogenized using the adaptive learning approach. Another empirical phenomenon that is well documented in the empirical literature is the decline in shock volatility, in a certain context referred to as the great moderation. If learning dynamics contribute to the model while at the same time fading out over time due to E–stability, this channel will initially inject additional variance into macro variables that will eventually clear out. Without learning taken into account the related reduction in variance must be traced back to a reduction in the variances of the shock term, leading to the interpretation of economic disturbances having become more favorable without being able to explain this phenomenon.

Surprisingly, though, there has not been much research devoted to the question of empirical relevance. After all, adaptive learning would matter little if its quantitative importance would be limited. The goal of this thesis is to provide empirical evidence on this approach and to demonstrate its power in accounting for otherwise hardly explainable phenomena.

There are two channels that inject additional dynamics into an economic system under adaptive learning. The first is the effect of frequent model revisions on optimal policy decisions. As soon as new information arrives the policymaker will reconsider the way he uses his instruments to influence his target variables. With a new transition law the optimal trade off between variables changes due to modified dynamic relations. This effect has recently been investigated by Primiceri (2006) in a model of US inflation and unemployment dynamics to study the impact of learning on the federal funds rate. Assuming that policy was conducted optimally he shows that the inflationary outbreak as experienced in the US during the great inflation episode during the seventies and early eighties can be traced back to learning related dynamics.

In chapter 1 we apply a similar method to both US and UK data and pursue the question which interest rates an optimizing policymaker would have set if he had been subject to imperfect information. We combine Primiceri's investigation of the Great Inflation using the adaptive learning approach with Orphanides' work on the importance of mismeasurements in unobservable variables such as output or unemployment gaps. We depart from the standard single equation approach to model the latent natural rate of unemployment and implement an appropriate signal extraction method to determine it. The advantage of this procedure is that it yields real–time estimates of the natural rate that are in line with narrative evidence and, unlike the standard method, can be equally applied to different data sets. We challenge Primiceri's assumption that policy was conducted optimally, which appears questionable at least in face of the tremendous outbreak of inflation in the United Kingdom in the seventies. Instead we compute the interest rate paths optimizing policymakers would have set had they been subject to imperfect information, and contrast them with historical rates. We find that historical and optimally recommended rates are highly correlated, but that during the great inflation historical rates were set moderately relative to what optimality considerations under certainty equivalence would have implied.

Chapter 2 picks up this idea and brings Brainard's (1967) uncertainty result into the analysis. Our primary question is whether outcomes as experienced in the US and the UK during the great inflation can be attributed to learning dynamics. We simulate inflation–unemployment dynamics as estimated from time series data given an optimizing policymaker subject to imperfect information. Again, we do not focus on a single type of policymaker, as characterized by a set of preference parameters, but investigate which outcome would have been experienced by different preferences. We discuss that a combination of wrong perceptions of inflation persistence and the Phillips curve slope, that measures the extend to which real activity impacts on changes in the price level, together with a substantial measurement error regarding the natural level of unemployment inevitably pushed the economy into a high inflation era. Our simulations demonstrate that this is the most likely outcome, with probabilities of it occurring almost reaching a hundred percent, while on the contrary without learning these high inflation episodes would rarely occur, and even if they did, they would not last as long and would not reach the high levels that we observe in the data. In fact, our learning model replicates the stylized facts encountered in the data, including the outbreak of high inflation, the fact that the peaks of the unemployment gap lag those of inflation a few quarters, and the rapid disinflation. We also find that the characteristics of the aforementioned combination of beliefs evolve quite naturally in our learning models. Despite our simulations generating episodes comparable to the great inflation, the extremely high inflation rates–in particular in the UK–are rarely attained in our simulations. Following the idea that the time varying uncertainty surrounding consecutive estimates might be the source of this behavior, we conduct a counterfactual exercise in the spirit of Brainard. We find that a learning policymaker who obliges himself to a mute response in face of high uncertainty would induce inflation episodes matching those historically observed in size and duration.

The second channel received more attention in the literature and relates to the impact of learning on expectations and the related feedback. Given the immense interest in the estimation of dynamic stochastic equilibrium (DSGE) models, chapter 3 develops a method that allows researchers to replace the assumption of rational expectations with that of real–time expectations. The procedure allows to estimate general DSGE models under the alternative assumption of adaptive learning and thus contributes to the literature in that it allows to keep track of the dynamic interaction of actual dynamic equations and the perceived law of motion. We apply this method in Ireland's (2004) variant of the New Keynesian model that assesses the role of real–business type technology shocks. Interestingly, we find that the estimation of his model under adaptive learning not only gives a better fit to the data, but also yields substantially different results. This shows that ignoring learning dynamics in the estimation introduces a substantial bias in the remaining estimates.

CHAPTER 1

Optimal Monetary Policy under Learning: An Empirical Approach

1.1. Introduction

Most modern macroeconomic models are characterized by forward looking behavior. The standard assumption on expectations formation is the rational expectations hypothesis, requiring agents to build forecasts consistently within the economic model. While this is a very satisfying assumption in view of microfounded optimizing behavior it is often criticized for its strong informational assumptions. Apart from the behavioral equations, agents must have perfect knowledge of aggregate dynamics which is in conflict with econometric practice. Adaptive learning (AL), advanced by Evans and Honkapohja (2001), is a recent approach to attenuate this assumption in that it assumes imperfect information and requires agents to infer precise knowledge of macroeconomic relations by real–time data observation.

In a New Keynesian model agents not only behave optimally in making their allocation and pricing decisions but also in the way they form their expectations. Many studies question whether the policymaker behaved optimally in that framework during the 1970s and find that a change in policy preferences or a change in the way policy is conducted led to the low inflation environment that we experience since the successful completion of the Volcker disinflation, see e.g. Clarida et al. (2000).

The adaptive learning approach attributes the bad performance to additional dynamics due to imperfect knowledge and the evolution of decision makers' beliefs. Since the Great Inflation was present not only in the US, it appears debatable whether the change in the Fed's chairmanship in 1979 could be the cause of the improved performance. Thus AL offers a unified approach to answer this question not only for the US but also for other countries with a qualitatively similar inflation history such as the UK.

Primiceri (2006) applies this methodology and investigates whether the high inflation episode of the 1970s and early 1980s is consistent with an optimizing policymaker who is exposed to imperfect information, and due to his restricted information set brings inflation to very high levels and for a prolonged period despite his best efforts. In particular he shows that the inflationary outbreaks would have equally well occurred under an optimizing policy regime that is exposed to imperfect information about the economic structure.

Primiceri's method starts from the conjecture that US monetary policy was at least close to optimal behavior under learning, but this assumption is not examined explicitly. It is hard to verify whether the policy was conducted non–optimally when the opposite is assumed from the outset. Although Primiceri demonstrates that a Great Inflation is possible even in case the central bank behaves optimally, his approach clearly cannot answer this question. Moreover, given the extremely high inflation in the UK in that episode it appears questionable whether this is indeed consistent with historical rates. Hence in this paper we challenge this assumption and fill this gap by investigating how a learning policy maker would have optimally conducted policy before, during, and following the Great Inflation under imperfect information.

We follow Primiceri and analyze optimal policy in a small empirical macro model similar in spirit to Woodford's (2003) New Keynesian model, adjusted to fit it to US and UK time series. The policymaker is assumed to face a standard quadratic loss function allowing for an inflation bias, but he has to infer information about the model's dynamics by real time data observation. While Primiceri investigates a single type of policymaker characterized by preferences that most closely describe historical rates in terms of optimal outcomes, we will derive the paths for nominal interest rates for a broad range of policy preferences and compare the resulting interest rate decisions with historical rates.

We will investigate whether policy as pursued in the last 20 years that is commonly classified as appropriate differs from previous policy behavior or whether the same outcomes would result even without changes in preferences. We will see that changes in preferences play only a minor role and are not the source of the superior inflation outcomes of the last two decades. We find a dominant role of imperfect information, in particular in the mismeasurement of the non–accelerating inflation rate of unemployment, as also documented by Orphanides and van Norden (2002), but also in the perception of inflation persistence and the Phillips curve slope as well as the uncertainty surrounding the corresponding estimates, as suggested by Brainard's (1967) study.

The paper is organized as follows. The next section reviews the historical performance of inflation and unemployment in the US and the UK, with particular focus on the Great Inflation episodes. Section 1.3 introduces the model and the adaptive learning approach while section 1.4 is devoted to optimal policy behavior in this context and discusses commonly assumed policy preferences underlying optimal decision making. Section 1.5 presents our results on interest paths of an optimizing learning central banker and contrasts his decisions with historical interest rates. Section 1.6 investigates the robustness of our findings to variations in the benchmark setting, and finally section 1.7 concludes.

1.2. Stylized Facts

The term Great Inflation (GI) refers to the prolonged high inflation period experienced in the United States, and similarly in other industrialized countries, for almost twenty years. Beginning in the mid–sixties, US inflation gradually increased, peaking at double digit levels in the mid–seventies and remained high for a sustained period, until it finally returned to low levels under Federal Reserve chairman Paul Volcker. This rapid decline in the early eighties is now referred to as Volcker Disinflation. A large amount of research has been conducted to explain this episode, but until now it seems difficult to explain all relevant empirical facts consistently in a model. This introduction gives a brief description of these facts and reviews some representative explanations that have been advanced in the literature.

Figure 1.1 (top panel) plots annualized quarterly inflation and the unemployment rate for US data. Beginning in the 1960s, inflation gradually increased, reaching a maximum of 12 percent in 1974. Although it decreased for a while, it remained high on average and peaked again above 10 percent by the end of 1980. This second peak was followed by a sharp disinflation, which quickly brought the inflation rate below 4 percent within two years and to an

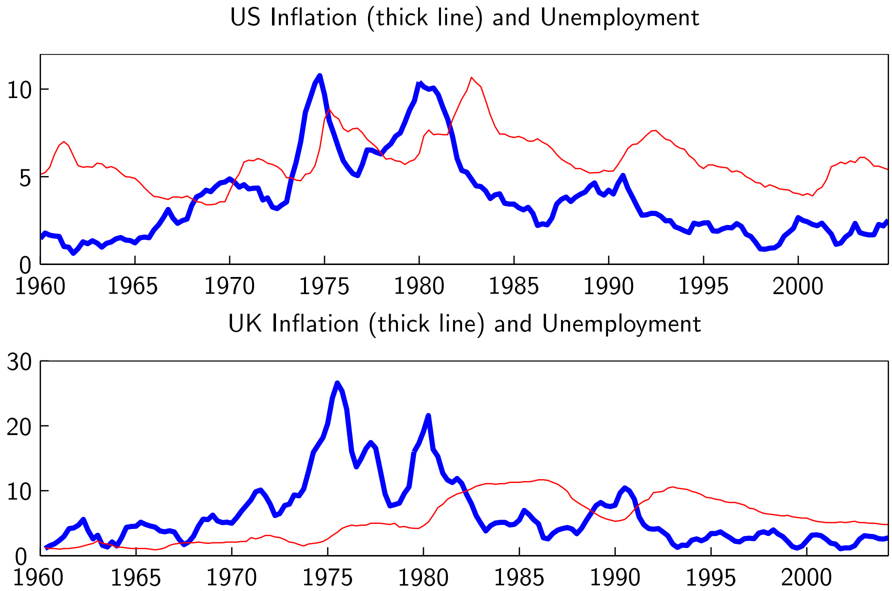

Figure 1.1: US (top panel) and UK (bottom panel) time series of inflation (thick line) and unemployment.

average level of 2.5 percent thereafter. Unemployment lags inflation, peaking one year and two years, respectively, after inflation.

Inspection of figure 1.1 (bottom panel) reveals that the situation in the United Kingdom was qualitatively similar to the US case discussed above, though it was far more severe. We observe slowly but continuously increasing inflation in the beginning of the sample, until inflation peaked at 26.6 percent in the second quarter of 1975. Although the economy briefly recovered, inflation remained high, peaked again five years later at 22 percent, until it finally dropped rapidly to levels below five percent, where it remained afterwards, except for 1987 and the following years where it temporarily touched eight percent. This rapid decline was followed by a prolonged period of unemployment that ceased only several quarters until inflation reached low levels. We also observe this phenomenon of a disinflation leading unemployment peaks in the periods following the inflationary years 1987–1992. To summarize, we observe UK inflation continually rising until it reached its peak level in the mid–seventies, and remaining at high levels for several years. The UK economy experienced a rapid disinflation in the early eighties, which was followed by an upshot of unemployment. Hence, the stylized facts discussed for the US were experienced in the UK economy as well.

1.3. The Model and Adaptive Learning

The adaptive learning literature typically analyzes related questions in a version of Woodford's (2003) DSGE model, see e.g. Evans and Honkapohja (2001) or Bullard and Mitra (2002) for prominent examples. This model consists of the New–Keynesian Phillips curve

$$\pi_t = \alpha_f E_t \pi_{t+1} + \theta x_{t-1} + \varepsilon_t \tag{1.1}$$

and an IS type aggregate demand equation relating a measure of real activity x to its future expected values, a policy instrument r, and a shock term η_t,

$$x_t = E_t x_{t+1} + \lambda r_t + \eta_t \tag{1.2}$$

Equation (1.1) is a pricing equation based on the Calvo (1983) model. In this model a random fraction ω of price setters is allowed to reoptimize their individual prices. The variable x is identified with real marginal costs, but in empirical work it is commonly approximated by either the output gap, e.g. Smets (2002), that is the deviation of real output from its flexible price benchmark, or by the unemployment gap, e.g. Primiceri (2006), the deviation of current unemployment from the rate which puts no pressure on the inflation rate. Orphanides and Williams (2007) emphasize that the unemployment gap x_t and the output gap x_t^o are related by a proportional relation known as Okun's (1962) law.

Equation (1.2) is a log–linearized Euler equation that results from agents' utility maximizing behavior in a representative agent model with monopolistic competition in the goods market. r is the ex ante real interest rate, relating to the central bank's policy instrument i, the nominal interest rate, via the Fisher equation $r_t = i_t - E_t \pi_{t+1}$.

All coefficients are functions of deeper behavioral or preference parameters, though we shall not be interested in this relation. What is important for our approach is the dynamic law that the central banker perceives when thinking about an appropriate path for his policy instrument.

Typically these microfounded equations are modified by including additional lag terms of the endogenous variables, to capture dynamics present in the data but not accounted for by the stylized model (1.1) and (1.2); see Fuhrer and Moore (1995). If the data favors the more restrictive theoretical model, these parameters would be estimated to be zero, thus nesting the underlying model. Ireland (2004) proposes a similar procedure. There is currently a debate in the adaptive learning literature, whether empirically measured inflation persistence is structural, as suggested by Fuhrer and Moore's hybrid specification with non–zero parameter estimates of the backward components, or whether learning dynamics account for at least part of it, making the backward looking term redundant. Thus, the burden to explain persistence is shifted from structural factors to learning related dynamics, see Milani (2005) and Junker (2008).

For our analysis of optimal monetary policy we prefer a reduced form system which is more in line with empirical macro models, given by

$$\pi_t = \alpha(L)\pi_{t-1} + \theta(L)x_t + \varepsilon_t^\pi \tag{1.3}$$

$$x_t = \rho(L)x_{t-1} + \lambda r_{t-1} + \varepsilon_t^x \tag{1.4}$$

where we use a lag order of two for the autoregressive polynomials $\alpha(\cdot)$, $\theta(\cdot)$ and $\rho(\cdot)$ thus obtaining a model identical to that used in Primiceri (2006) and similar to Rudebusch and Svensson's (1998) and Smets's (2002) models. Following Rudebusch and Svensson and Smets we compute the real rate as $r_t = i_t - \bar{\pi}_t$, where our measure of expected inflation is a four quarter moving average of current and past inflation rates.

The timing assumption in these models, as in our equations (1.3) and (1.4), implies that monetary policy affects real activity with a one period lag and the unemployment gap in turn has an impact on inflation one quarter ahead. Overall, monetary policy affects inflation with a two quarter lag. This is consistent with the dynamics of the monetary transmission mechanism as documented by Christiano et al. (1999), and allows us to solve for optimal monetary policy using standard methods, as explained below.

The advantage of using an empirical reduced form specification is that we can focus on the behavior of a learning policymaker while suppressing the expectations channel. Such a reduced form can be obtained by modelling expectations as sufficiently backward looking, adaptive processes. Models that emphasize the connection between structural and reduced form are subject to chapter 3 and Junker (2008). It would be an ambitious goal to analyze the structural model, using two–sided learning. This is left for future research.

It should be noted, however, that two–sided learning is substantially more complex and is rarely applied in the AL literature, mostly in calibrated models of the economy. We interpret our specification as a tractable reduced form version of the standard DSGE model.

Perceived Law of Motion. We equip the policymaker with a dynamic model of inflation and unemployment of the form discussed previously, so that he faces the equations

$$\pi_t = \hat{c}_\pi + \hat{\alpha}_1 \pi_{t-1} + \hat{\alpha}_2 \pi_{t-2} + \hat{\theta}_1 x_{t|t-1} + \hat{\theta}_2 x_{t|t-2} + \varepsilon_t^\pi \tag{1.5}$$

$$x_t = \hat{c}_x + \hat{\rho}_1 x_{t|t-1} + \hat{\rho}_2 x_{t|t-2} + \lambda(i_{t-1} - \bar{\pi}_{t-1}) + \varepsilon_t^x \tag{1.6}$$

where constants have been added so that the model can be estimated using historical data series, in contrast to the theoretical model which is formulated in terms of deviations from steady state values which are usually unknown.

This specification is referred to as perceived law of motion, since it specifies the model underlying the policymaker's decisions given his perception of the economic environment, which does not necessarily coincide with the true data generating process. Our main assumption is that the central banker has imperfect information about the model's dynamics and he will use statistical inference to update his beliefs, which will crucially drive his optimality decisions.

The key assumption of adaptive learning is imperfect knowledge of decision makers. The central banker is assumed to know the functional form of equations (1.5) and (1.6), but he lacks information regarding the parameter values. He thus uses statistical inference to improve his estimates of the autoregressive parameters, the Phillips Curve slope parameters, and the constants over time, which we collect in the belief vector

$$\beta_t = (\hat{c}_{\pi,t}, \hat{\alpha}_{1,t}, \hat{\alpha}_{2,t}, \hat{\theta}_{1,t}, \hat{\theta}_{2,t}, c_{x,s}, \hat{\rho}_{1,t}, \hat{\rho}_{2,t})' \tag{1.7}$$

where the index t refers to the information set upon which the estimator is based. The date–t information set includes $I_t = \{\pi_s, u_s, r_{s-1}\}_{s \leq t}$, which implicitly includes the unemployment gap since it is a function of observable variables included in the information set.

The impact coefficient λ is difficult to estimate over the whole set of subsamples. We will see below, when discussing the standard deviations of the estimates over time, that during particular periods the precision of estimates is rather poor; in particular the estimates of λ are insignificant for many subsamples and even display the wrong sign occasionally. We believe that it is plausible that central bankers at least agree on the sign of this parameter and prefer to calibrate this coefficient on a standard value.

It is a rather common phenomenon in the adaptive learning literature that the same parsimonious model is not capable of giving a satisfactory fit for all subsamples, and for this reason, many applications focus on simple models, allowing learning of only one or a small subset of parameters of central interest while calibrating the remaining coefficients. We will nevertheless estimate all dynamic parameters, as indicated by the belief vector β.

Primiceri also implicitly fixed this parameter on a value $\lambda = 0.024$, while other studies find somewhat larger values, albeit usually relating to the output gap. If we follow Orphanides and Williams's suggestion in choosing an Okun coefficient of (minus) 2, we can translate estimated or calibrated impact coefficients from an output gap equation into corresponding impact coefficient for our unemployment gap equation (1.6). In this regard, we find values ranging from $\lambda = 0.05$ (cf. Rudebusch and Svensson (1998)), $\lambda = 0.025$ and $\lambda = 0.03$ (for different subsamples in Smets (2002)). We will also consider somewhat larger values and extend the range to allow for values above Rudebusch and Svensson's estimate. Consequently, we will consider $\lambda \in [0.024, 0.08]$ as the relevant range, choose the midpoint as benchmark value, which corresponds to Rudebusch and Svensson's estimate, and consider variations in the full range in a separate sensitivity analysis. Our robustness check confirms that the precise value of the impact coefficient does not affect our results qualitatively.

Nairu as a latent variable. Our model features a time–varying rate of unemployment which is thought of as the level that exerts no pressure on the inflation rate. With a slight abuse of (technical) language, this particular rate is commonly referred to as the non–accelerating inflation rate of unemployment, or Nairu for short, since it is the rate that induces no movement of inflation, cf. e.g. Ball and Mankiw (2002) or Gordon (1997). The unemployment gap is the difference between the current unemployment rate and the prevailing Nairu, hence the latter is formally given by the level of unemployment for which the unemployment gap in equation (1.3) vanishes. Importantly, in the present context, at date s the policymaker also updates the whole path of his perceived Nairu given a new data observation, yielding the series $\{u^*_{t|s}\}_{s\leq t}$ where $u^*_{t|s}$ is the date–t estimate of the Nairu level prevailing at date s. So before we can proceed with describing the policymakers learning algorithm, we have to discuss how he infers the latent variable x from the data.

Note that the gap is determined by contrasting the level of the corresponding date's unemployment rate to the Nairu as perceived by the policymaker at the current date t. Since this is also an estimate, it generally differs from its true level, hence the policymaker does not only face uncertainty about the model's true parameter values, but is also subject to misperceptions of one of its main target variables. This last point is the crucial feature in Orphanides' work. He analyzes the effect of real time and 'quasi–real time' data, the latter being the effect present in our approach; in Orphanides and van Norden (2002) both effects are shown to be substantial.

At this point it is worth discussing how related papers handle the issue of Nairu estimation in a learning environment. Orphanides and Williams (2007) note that given the time variation in the Nairu policymakers need to continuously reestimate this variable in quasi–real time. As a simplifying approach he suggests the use of a simple algorithm which basically extracts the Nairu as a recursive sample mean of unemployment, though the application of a constant gain approach additionally involves a discounting of past observations (see the related discussion below). Primiceri (2006) adopts this procedure and calibrates the learning gain such that the resulting series is broadly consistent with conventional wisdom regarding the Nairu path. This procedure works approximately for US data where the Nairu is regarded as a very smooth series with small fluctuations around the mean unemployment level of six percent. Applying this method to UK data would produce a Nairu path that continually underestimates the path relative to conventional empirical estimates which display higher fluctuations than the US counterpart. To avoid this problem we will employ a more sophisticated method that is designed to extract a unit root latent variable from observable data.

Standard methods to extract trends are Kalman's (1960) Filter (cf. Gordon (1997) for an application to US data, Franz (2005) for German data and Batini and Greenslade (2006) for UK data, or for the case where the excess demand term is measured by the output gap, see Smets (2002)), and the Hodrick–Prescott Filter (see Hodrick and Prescott (1997); henceforth HP). The Kalman Filter simultaneously estimates the path of the Nairu and the parameters of the model. Application of the HP Filter proceeds in two steps. The filter is appropriate to identify the unit root trend component of unemployment, and the resulting difference to the unemployment series is subsequently used as a regressor in either ordinary least squares regression or a commonly used version of it as discussed below.

For simulations in connection with adaptive learning it is advisable to choose Hodrick and Prescott's method since it is substantially faster and, more importantly, numerically more reliable than the Kalman Filter approach.[1] However, both methods yield very similar results. In fact, the arbitrary choice of the HP–smoothness parameter is made as to assure that the resulting series for the Nairu is in line with previous studies, as will be discussed below. The choice of this parameter and the robustness of the results against changes in this parameter are the subject of section 1.6. For an application of the Kalman Filter to extract the gap variable, see chapter 3 and Junker (2008).

Learning Algorithms. After deriving u_t^*, the policymaker computes the perceived unemployment gap $x_t = u_t - u_t^*$ and applies appropriate estimation techniques to the pricing and demand equation, yielding a revised estimate β_t.

A common approach in the learning literature, suggested for example by Evans and Honkapohja (2001) and applied in important contributions, e.g. in Bullard and Eusepi (2005) or Cho et al. (2002), are the so–called constant gain learning algorithms. They are favored over related methods such as ordinary least squares since its discounting of past data reflects the desire of agents to keep track of regime breaks. In particular in face of models under learning, where feedback effects usually affect the dynamic evolution this appears as a plausible procedure.

Related work, such as those cited above, typically present a recursive formulation of the learning algorithm, which integrates the last observation into

[1]Maximum likelihood estimation using the Kalman Filter requires numerical optimization of the same model for an increasing sequence of data points; usually additional assumptions regarding initialization or variances are made depending on the outcome of the filter. This procedure is not advisable for simulations.

the previous estimate. This formulation is convenient since it offers a computationally fast updating procedure for a growing sample. However, in face of latent variables this procedure in inappropriate. Each time the model is reestimated for an increased sample the whole regressor based on the unobservable component will change, so that the recursive procedure would ignore the improved estimate of the path except for the most recent realization. Thus it should be noted that recursive and ordinary least squares are not equivalent once we determine a latent variable in a first step.

It is nevertheless illustrative to consider the standard recursive learning algorithm for a single equation regression of y on x,

$$\beta_t = \beta_{t-1} + \gamma_t \Sigma_t^{-1} x_t (y_t - x_t' \beta_{t-1}) \tag{1.8}$$

$$\Sigma_t = \Sigma_{t-1} + \gamma_t (x_t x_t' - \Sigma_{t-1}) \tag{1.9}$$

where Σ_t is the date–t estimate of the covariance matrix of the OLS estimator. The estimate β_t for a sample ending in period t can be computed from the previous estimate from the sample ending in $t-1$ and the most recent observations (y_t, x_t). The estimate is adjusted by a weighted forecast error from the previous regression. The factor γ_t, referred to as gain parameter, plays a key role: for $\gamma_t = 1/t$ we have a recursive formulation of ordinary least squares where the same weight is attached to each observation. Constant gain algorithms differ from least squares in replacing this time varying gain by a constant gain $\gamma_t = \gamma$ for all t. Further details are provided in appendix A.1.

To be able to use a constant gain algorithm even in face of the latent unemployment gap and thus enabling us to relate our results to similar studies using particular values for the gain parameter, we will use discounted least squares (DLS). As also discussed in appendix A.1, this is the non recursive counterpart to constant gain algorithms, with gain and the corresponding discount factor being related as

$$\gamma = \sqrt{1 - \delta^2} \tag{1.10}$$

In this formulation a zero gain corresponds to the borderline case of a unit discount factor, in which case DLS coincides with ordinary least squares (OLS), so using gain values within a standard range while explicitly allowing for the case $\gamma = 0$ nests OLS as a special case. As an example, the widespread calibration for the gain value $\gamma = 0.03$ is associated with a discount factor of $\delta = 0.985$.

At date t the policymaker will thus proceed as follows: firstly, he revises his estimated Nairu path, $\{u^*_{t|s}\}_{s \leq t}$, and deduces the perceived unemployment gap for all dates, $\{x_{t|s}\}_{s \leq t}$. If he wants to use a gain such as our benchmark value γ_0 he uses equation (1.10) to solve for the corresponding discount factor δ, discounts all regressors and endogenous variables with this factor and applies OLS to the transformed system. For $\delta < 1$ this method is DLS. The parameter estimates derived in this way will then be mapped onto an appropriate state space form to revise his optimality problem within his updated dynamic transition law of the economy, as discussed in the next section.

Learning results. The following figures track the evolution of the policymaker's beliefs of the corresponding parameters over time. While all samples start in 1955 and 1959, resp., depending on data availability, the time axis depicts the final sample point which is also the estimation date, e.g. in figure 1.2, the value of approximately 0.5 in 1964 indicates that inflation persistence was estimated at that level at that time while 16 years later, in 1980, inflation was perceived to be substantially less stationary with a value close to unity.

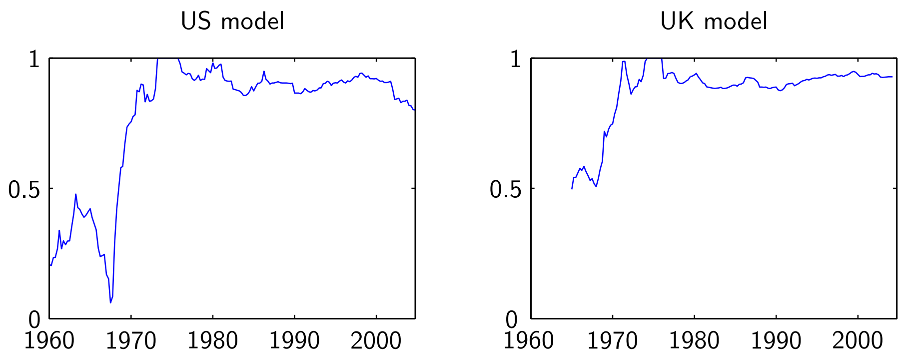

Figure 1.2: Beliefs of Inflation Persistence

Beliefs of Inflation Persistence. Both the evolution of beliefs about US and about UK inflation persistence share close similarities. Initially policymakers regard inflation as a strongly mean reverting process with an autoregressive parameter of around 0.3 and 0.5 implying quick return to the target level after occurrence of shocks. This view drastically changes in the seventies when updated parameter estimates indicate that inflation is in fact close to a unit root process.

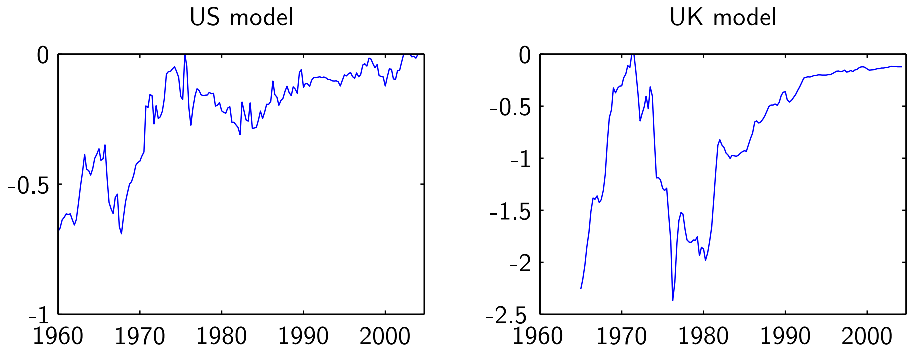

Figure 1.3: Beliefs of the Phillips curve slope

Beliefs of the Phillips curve slope. The evolution of the beliefs about the Phillips curve slope are similar in both models. Initially, the slope has a substantial, negative value that is continually eroded until it reaches zero in the early seventies. This means that the sacrifice ratio, that is the units of unemployment above the Nairu necessary to induce a desired effect on inflation is bit by bit increased. After the effectiveness almost completely vanished, the perceived slope improves in favor of active policy making, taking on higher and higher values (in absolute terms) until it reaches a level of maximum efficiency around 1980, somewhat later in the US, and a few quarters earlier in the UK. Afterwards, the slope coefficient slowly approaches values close to zero. This observation is consistent with the findings in Primiceri (2006). Using different price measures, other studies find even more pronounced changes in perceived persistence though they agree qualitatively with our results, cf. Milani (2005).

The amplitude of the UK slope perceptions is higher since the higher assumed time variability of the Nairu soaks up a great part of the unemployment fluctuations resulting in a lower gap amplitude. Ceteris paribus, a lower amplitude of the estimated unemployment gap yields higher parameter estimates for the slope. It is the product of these two which can be inferred from the data, the separation in gap and slope multiplier depends on the assumptions regarding the Nairu smoothness.

Full sample and quasi–real time Nairu. Figure 1.4 plots the unemployment rate, the full sample Nairu estimate (dashed line) and the quasi–real time

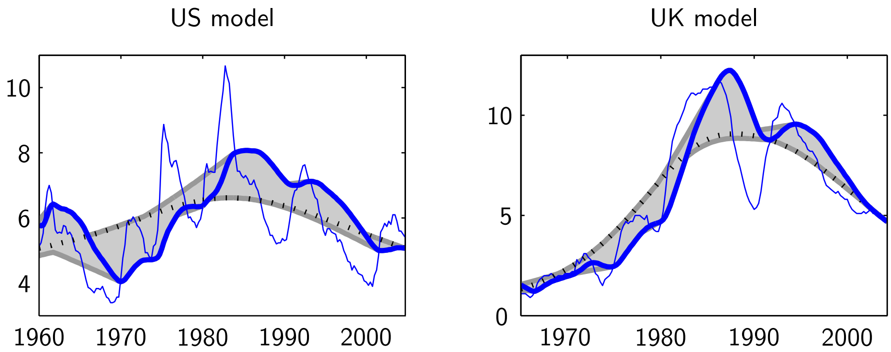

Figure 1.4: Full sample Nairu (dashed line), quasi–real time Nairu (thick line) and revision intervals (shaded region)

Nairu. The latter is defined pointwise, its date t value giving the perception of the Nairu as of time t, i.e. formally $u^*_{t|t}$, whereas in contrast the full sample Nairu plots the date T perception of the Nairu level at any date within the sample, i.e. its date t value is $u^*_{T|t}$. The concept of a (quasi–)real time Nairu is frequently employed by Orphanides, e.g. in Orphanides and Williams (2005). Comparing our US quasi–real time Nairu with the corresponding one in the latter paper, we find that both paths are very similar, although our Nairu has a higher amplitude in the years after 1981. The full sample Nairu is a smooth series and is comparable to results in other empirical studies, cf. Gordon (1997) for an example. The difference between both series gives the real time misperception of the latest Nairu values, that is in the early 1970s the perceived values of the Nairu prevailing at that time was slightly above 4 percent while current estimates for these values are substantially higher, being around 6 percent as is also implied by our full sample estimates.

1.4. Optimal Policy under Adaptive Learning

We will impose a quadratic loss function for the central banker which apart from the incorporation of a potential inflation bias is standard in the literature. Optimal monetary policy will be a feedback rule that is derived from this underlying loss function. We will discuss the policymakers optimization problem, investigate the effect of adaptive learning on these decision and consider the relevant range for the preference parameters in the final part of this section.

The Optimal Linear Regulator Problem. At any date t the policy-maker chooses a path for its instrument v to minimize the discounted stream $\sum_{s\geq t} \delta^{s-t}\mathcal{L}_s$ of current and expected future period losses, which are given by

$$\mathcal{L}_s = (\pi_s - \pi^*)^2 + \omega_u(u_s - k\hat{u}^*_{t|s})^2 + \omega_\nu(v_s - v_{s-1})^2, \quad s = t, t+1, \ldots \tag{1.11}$$

The loss function is quadratic in deviations of inflation π from a target rate π^*, in the unemployment gap $u - u^*$ and in the rate of change of the policy instrument. The standard quadratic term in the unemployment gap is modified to allow for an inflationary bias of the central banker towards the real side of the economy, that is towards unemployment. Instead of subtracting the Nairu from the level of unemployment, we subtract $\kappa\hat{u}^*_t$ to obtain the standard case for $\kappa = 1$, and an increasing tendency to allow for an inflationary bias in favor of unemployment rates being below the Nairu as κ approaches zero.

The change in the policy instrument is taken account of to reflect the degree of interest rate inertia typically found in empirical studies, which is not implied by the model. However, studies usually find an extremely large response of interest rates towards deviations of inflation or unemployment from their respective target values, unless the smoothing objective is explicitly taken account of. We normalize the weight on inflation deviations to one, so that the parameters ω_u and ω_ν are relative weights.

The policy maker faces a dynamic transition law given by equations (1.5) and (1.6), which is linear in the state vector $y_t = (1, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, u_t, u^*_t, u_{t-1}, u^*_{t-1}, v_{t-1})'$. The policymaker's dynamic program can be reformulated as a linear–quadratic optimal regulator problem of choosing a path for the policy instrument v that minimizes the quadratic loss function

$$\sum_{t\geq s} \delta^{t-s}[y_t'\Omega_{yy}y_t + v_t'\Omega_{vv}v_t + 2y_t'\Omega_{yv}v_t] \equiv \sum_{t\geq s} \delta^{t-s}(y_t', v_t)\Omega(y_t', v_t)' \tag{1.12}$$

subject to the perceived linear transition law

$$y_{t+1} = Ay_t + Bv_t + \varepsilon_{t+1}, \tag{1.13}$$

with a singular weight matrix Ω, and system matrices A and B. The singularity of the matrix can be addressed by the invariant subspace method using the generalized Schur decomposition, as discussed e.g. in Hansen and Sargent (2005). The procedure requires transformations to eliminate discounting and the presence of mixed terms. Discussion of this method along with details on the matrices are deferred to appendices A.3 and A.2. The resulting optimal policy depends on the policymakers preference parameter p to be discussed in detail below, which comprises his discount factor, the inflation target, the inflationary bias parameter and the weights in the loss function. The solution can be expressed as a linear feedback rule

$$v_t = -Fy_t \tag{1.14}$$

where the elements of the row vector F are computed numerically using the procedure proposed by Hansen and Sargent.

Combining Learning and Optimal Policy. Inserting an optimal policy reaction function as described in equation (1.14) into equation (1.13) yields the implied optimal dynamic behavior of the system, given by

$$y_t = (A - BF)y_{t-1} + \varepsilon_t \tag{1.15}$$

that the policymaker expects in the absence of shocks. As is standard, the corresponding path for the instrument will only be implemented in expectation. Upon arrival of new information, i.e. the realization of shocks, this path will be adjusted. The feedback rule on the other hand usually stays invariant and is designed to accommodate the impact of shocks. In particular, this means that the system evolves according to the same dynamics, in the sense of having the same eigenvalues and thus adjustment speeds, though from different initial values that depend on the shock realizations.

Under learning, however, also the feedback rule, given by the vector F changes each period. New data does not only include (at least approximate) observations of shocks as under RE, but it also supplies valuable information about the unknown parameters of the dynamic transition law, and consequently also calls for a readjustment of the rule itself. The policymaker applies his learning algorithm to equations (1.5) and (1.6) each quarter as new information arrives. This gives him an updated linear transition law of the form

$$y_{t+1} = A_t y_t + B_t v_t + \varepsilon_{t+1} \tag{1.16}$$

where the matrices A and B are now indexed by t, the time the estimation is conducted.

It has become standard in the literature on adaptive learning to follow Kreps (1998) in assuming what he calls Anticipated Utility. This stands for the assumption that the policymaker treats the dynamic system (1.16) as time invariant and solves the corresponding optimal regulator problem without accounting for potential future adjustments in the system matrices. We interpret this as follows: the underlying parameters are regarded as constant while fine–tuning gives improved estimates each period. Thus even if he is continually reassured that the parameters differ from period to period, the policymaker does not attribute these changes to a changes in the model but to more precise understanding of the latter. Cogley and Sargent (2001) revisit this assumption, provide evidence that this is a valid simplification in our environment and promote its use to keep the analysis tractable. We thus obtain a linear feedback rule responding to the state vector y,

$$v_t = -F(p, \beta_t) y_t \tag{1.17}$$

where the notation emphasizes the dependence of the feedback rule on both date–t beliefs β_t and the preference vector p that governs the decision problem.

To summarize, at any date t the policymaker with fixed preferences p observed data up to this date, his information set thus being $I_t = \{\pi_s, u_s, i_s\}_{s \le t}$, which implicitly includes unemployment gap since it is a function of observable variables included in the information set. He derives an estimated path of the unemployment gap and applies a non–recursive learning algorithm as discussed above yielding updated perceptions β_t. He solves the corresponding optimal regulator problem, obtaining an optimal path $\{-F(p, \beta_t)\ E_t y_s\}_{s \ge t}$, from which he implements the first prescribed policy move in the current period. Economic variables realize and the policymaker repeats these steps in the following period. Eventually, this yields a sequence of implemented optimal policy decisions, $\{-F(p, \beta_t) y_t\}_{t=1,\dots,T}$. It depends on the policymaker's preference parameters, that is his inflation target, the inflationary bias parameter and the weights in the loss function, collected in the preference vector p, which the next section discusses in detail.

Policy Preferences. We want to investigate how a standard policymaker will set policy in the economies described by our sequential estimates of US and UK dynamics. We thus pose the question of what interest rates would have been set at a given date t assuming that the policymaker behaved optimally given his most recent model estimates. Since this decision depends on the central banker's preferences we investigate a range of preferences that we consider relevant for policy making. We randomly draw a policy preference vector p using the uniform distribution over plausible intervals usually supported by the literature. Recalling that policy preferences are summarized by a vector $p = (\kappa, \omega_x, \omega_v)'$, we have to discuss plausible ranges for each of these parameters from which we will make uniformly random draws.

For the central banker's inflation target we choose a benchmark value of $\pi_0^* = 2$ which is a common value, cf. e.g. Bullard and Eusepi (2005) and Schorfheide (2005). Our main simulation in the next section allows for values in the range $\pi^* \in [1.5, 4]$.

We observe that the inflation bias parameter is constrained to the unit interval and predispose that very low values are not very likely. Given Primiceri's finding that the inflation bias is quantitatively negligible we consider the range $\kappa \in [0.5, 1]$ in our main part and choose his estimate $\kappa_0 = 0.87$ as benchmark.

Reasonable benchmark values for the relative weights in the loss function are $\omega_u = \omega_v = 1$ since this choice implies that all target values are equally important. It appears plausible, however, to allow each of the three variables to be a dominant target for monetary policy, so that relative weights above one should be regarded in the same way as their inverses, being below one. We choose a range of $[1/2, 2]$ for both weights implying that we consider central bankers with inflation aversion being four times stronger than one or both of the other targets, as well as the reverse, e.g. a disposition to interest rate smoothing being up to twice as high as either inflation or unemployment stabilization. Of course, since the parameters are jointly drawn from uniform distributions over the relevant range, we allow for all intermediate combinations as well, thus covering a broad range of relevant policy preferences.

1.5. Simulating Optimal Policy

1.5.1. Optimal reaction coefficients under adaptive learning. To simulate optimal policy decisions under learning we infer the successive estimates of the model as described in the section on adaptive learning, for each draw of the policy preference vector we compute the optimal feedback rule for each time t given the prevailing estimates as of that date, and evaluate the implied date t optimal interest rate using the latest estimate of the Nairu along with the observations of inflation and unemployment.

The optimal policy is a feedback rule linear in the 10–dimensional state vector, which as is clear from the above is time dependent. As an example, the rule for the benchmark US central banker in 1965 is given by

$$i_t = 0.3 + 0.1\pi_{t-1} + 0.08\pi_{t-2} + 0.03\pi_{t-3} + 0.01\pi_{t-4} - 0.6x_t + 0.8x_{t-1} + 0.7i_{t-1} \tag{1.18}$$

while the rule in 1990 is given by

$$i_t = 2.1 + 0.3\pi_{t-1} + 0.15\pi_{t-2} + 0.04\pi_{t-3} + 0.02\pi_{t-4} - 1.7x_t + 1.3x_{t-1} + 0.6i_{t-1} \tag{1.19}$$

Instead of reporting all ten coefficients and their evolution over time, we focus on useful summary statistics, the sum of the feedback coefficients on the lags of inflation, $G_\pi = F_{\pi_1} + \ldots + F_{\pi_4}$, the analogous sum corresponding to unemployment, $G_x = F_{u_1} + F_{u_2}$, and the inertial parameter, the coefficient on the lagged nominal interest rate, G_v.

Noting that the coefficients on u_t and u_t^* as well as those on their lags sum to zero[2], which means that although the response to the state vector allows for unrestricted responses to each unemployment and the Nairu, it in fact implies a feedback to the unemployment gap instead of its two components. We can compare the optimal policy rule with a Taylor rule by observing that a simple rule that approximates our optimal feedback rule is given by

$$i_t = G_c + G_\pi \pi_t + G_x x_t + G_v i_{t-1} \tag{1.20}$$

and recalling that the traditional Taylor rule for the target interest rate

$$i_t^* = \bar{r} + \pi^* + g_\pi(\pi_t - \pi^*) + g_x x_t \tag{1.21}$$

relates to the actual interest rate which also incorporates a smoothing objective via

$$i_t = \rho i_{t-1} + (1-\rho) i_t^* \tag{1.22}$$

[2]In the case $\kappa = 1$; otherwise the inflationary bias yields an extra term in unemployment, which we will ignore for this discussion.

we can map our more complex dynamically optimal rule on a Taylor type rule to obtain

$$g_\pi = G_\pi/(1 - G_v) \tag{1.23}$$

$$g_x = G_x/(1 - G_v) \tag{1.24}$$

$$\bar{r} = [G_c - (1 - G_v - G_\pi)\pi^*]/(1 - G_v) \tag{1.25}$$

This reduction to a simple rule facilitates its comparison to other studies, in

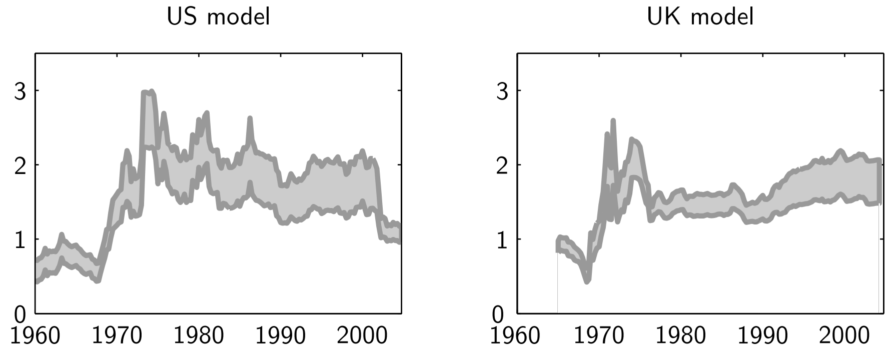

Figure 1.5: Approximate interest rate response to inflation. The shaded area represents the response coefficient for all preferences under consideration.

that our summary measures from the optimal feedback rules can be contrasted to standard Taylor rule coefficients. The evolution of the optimal feedback to inflation is plotted in figure 1.5, which offers shaded areas containing the optimal feedback coefficients for all parameterizations of policy preferences under consideration. A striking feature is that the optimal response to inflation endogenously starts rising from values below one in the early seventies and reaches its maximum just before the disinflation period begins, the time that is usually considered to mark the switch from bad policy to sound policy making that eventually ended the GI. This finding sheds new light on this debate, as it shows that a particular type of optimizing policymaker would have acted in the same way as is typically attributed to Fed chairmen before and after Paul Volcker. In our model the policymaker's preferences are fixed over the whole sample under consideration, in particular, he faces the same objective function before and after 1980, yet the stance against inflation substantially changes. In our model this is entirely due to the learning dynamics. The policy that is repeatedly identified as inappropriate in other studies (e.g. Clarida et al.) differs from the policy in the latter years only in the degree of imperfect information that the policymakers were facing when making decisions. Furthermore, US and UK policies converge to a similar behavior against inflation as in the latter part of the sample the feedback coefficients are almost identical. The feedback coefficients on the unemployment gap are plotted in figure

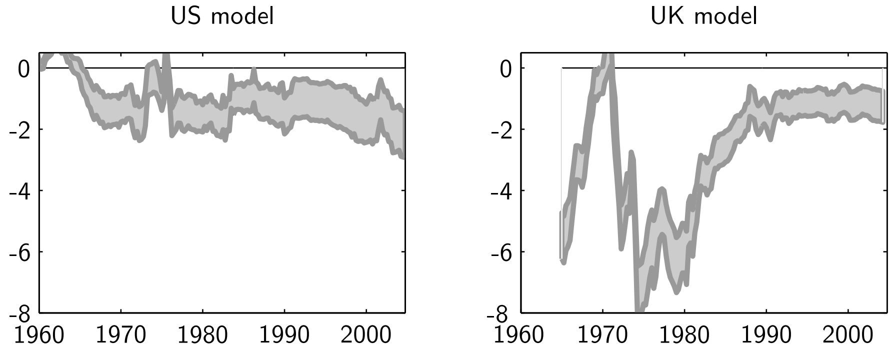

Figure 1.6: Approximate interest rate response to the unemployment gap. The shaded area represents the response coefficient for all preferences under consideration.

1.6. Apart from the large swings in the UK model which are due to the strong fluctuations in the perceived Phillips curve slope (cf. section 1.3), both models eventually approach a common value between −1 and −2, depending on the exact specification of preferences. Optimal interest rate inertia as depicted in

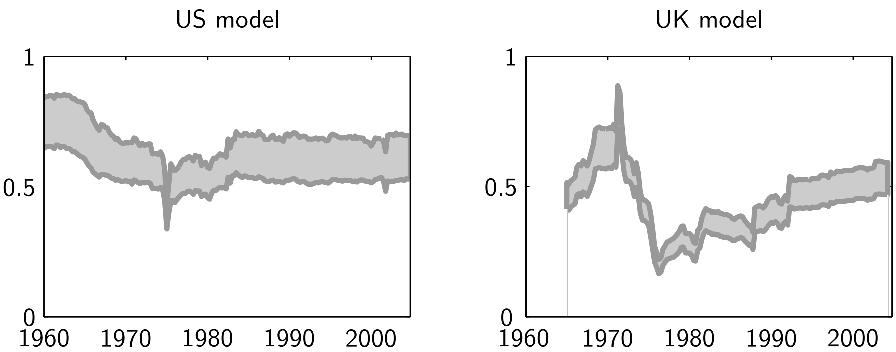

Figure 1.7: Interest rate smoothing for all preferences under consideration, summarized by the shaded area.

figure 1.7 has been quite stable in the US, falling slightly from initial values around 0.8 to somewhat lower levels around 0.6. In the UK model it had an upward tendency in the first quarters reaching a level of 0.9 in 1970 and after a subsequent drop to 0.3 in the first part of the seventies it slightly adjusted to values around 0.5 for the last 10 years of the sample. It should be noted that the smoothing component will be lower in the complex dynamic rule due to the additional channel from the lagged terms.

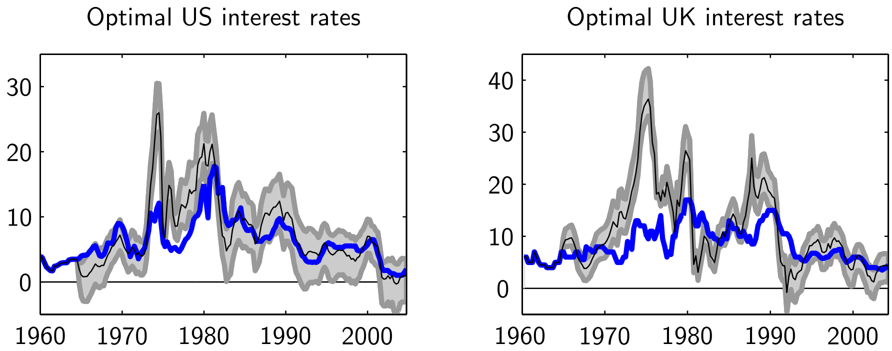

Figure 1.8: Simulated optimal interest rates for our benchmark calibration (line in the shaded region), for different preference parameters (shaded region) and historical interest rates (thick line).

1.5.2. Optimal policy paths under adaptive learning. Figure 1.8 summarizes our main results. It contains a range of optimal interest rate paths a learning policy maker would have set given his preferences can be expressed as a quadratic loss with parameters from the distributions discussed in the previous section. The figure also depicts actual interest rates as set by the Fed (thick line in the left panel), and the Bank of England (thick line in the right panel). The dashed line is the result of our benchmark calibration. There are several interesting observations to be made.

First, although we allow for a wide range of preferences, the resulting policies are qualitatively similar and highly correlated, thus policymakers with quite different preferences would have behaved rather similar in terms of characteristics such as maximum and minimum rates and turning points in the setting of interest rates.

Second, while historical interest rates and optimal ones were very close to each other–and almost coincide for the post–1980 subsample, they display substantial differences in the early part of the sample. This is true for the US case but in particular for the UK economy where both the deviations from

	US model		UK model	
	Correlation	Ratio	Correlation	Ratio
GI episode	0.73	0.77	0.50	0.59
Post GI episode	0.98	1.01	0.75	1.13
Full sample	0.86	0.94	0.58	0.98

Table 1.1: Correlation of optimal interest rate paths with historical rates and the ratio of both series.

optimality and eventually also the extend of the GI were more pronounced. We therefore conclude that historical monetary policy was not entirely consistent with optimal behavior during the Great Inflation period.

However, our third observation is the strong comovement of historical rates with interest rates set by an optimizing but learning central banker. Indeed, even during the Great Inflation period the corresponding correlations are very high, as is documented by table 1.1, although historical interest rates were persistently lower, with the ratio being 0.8 in the US model and 0.6 in the UK model. Chapter 2 will investigate this finding in more detail. The post–1980 period is characterized by optimal rates coinciding with historical ones, the ratio being approximately unity.

Finally, we note that qualitatively the results for the US economy are identical to those of the UK economy. In both the US and the UK case we find that the discrepancy between actual and optimal rates widens in the advent of the high inflation episode where an optimal regulator would recommend very aggressive interest rates to counter the rise in inflation, and observe this gap to close as strong interest rate movements–consistent with optimal policy rates–in both countries finally bring inflation back to moderate levels, where it is maintained until today with interest rate decisions being perfectly consistent with optimal behavior of a learning policy maker for the remaining 25 years of the sample.

1.6. Sensitivity Analysis

In the previous section we found that neither US nor UK interest rates were consistent with recommendations of a learning optimal regulator. However, one might argue that while we already considered a broad range of policy preferences, there might be specifications which are better capable of reproducing historical rates. Thus in this section we shall review optimal policy paths but reassess the effect of all relevant parameters. In each step we will fix all parameters but one on their benchmark values, enlarge the intervals of this free parameter to cover a sufficiently broad range, and investigate the effect of varying this parameter. Among them we will consider the HP filter smoothing parameter μ and the gain value γ of the adaptive learning algorithm, so we cover all relevant specifications of learning. Additionally, we will investigate the sensitivity of our results towards different values of the policy impact coefficient λ and we will investigate the contribution of varying single policy preference parameters. This will allow us to explore whether our main finding of non–optimality was an artifact of a too narrow specification or whether it holds more generally.

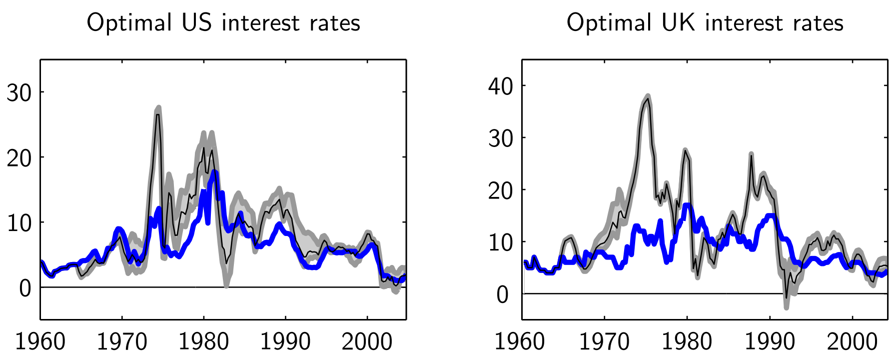

Figure 1.9: Optimal interest rates for varying weight on unemployment stabilization, $0.25 \leq \omega_x \leq 4$.

Policymakers with different weights on unemployment stabilization. At first, we shall be interested in varying the importance the policymaker attaches to unemployment stabilization, as expressed by the relative weight in the policy function, ω_u. We consider values that make the concern about inflation four times as large as that about unemployment stabilization, i.e. $\omega_u = 1/4$ and vice versa, as well as all intermediate values.

As is clear from figure 1.9, our main result still holds under all possible weight parameters. While changing the focus on unemployment yields paths for the US rate that differ in particular during the GI episode, all recommended rates remain substantially higher then historical rates. Again, preceding and following the GI historical rates are within the range of interest rates set by an optimizing learner, except for the period around 1990 where inflation rates reached the highest post–GI levels in our sample and where historical rates appear too conservative, again. In the UK case the changes in the relative weight on the unemployment gap has little impact on the resulting optimal interest rates. In particular in periods with extremely high inflation rates the unemployment objective appears to be subordinate, so that changing its weight has little effect.

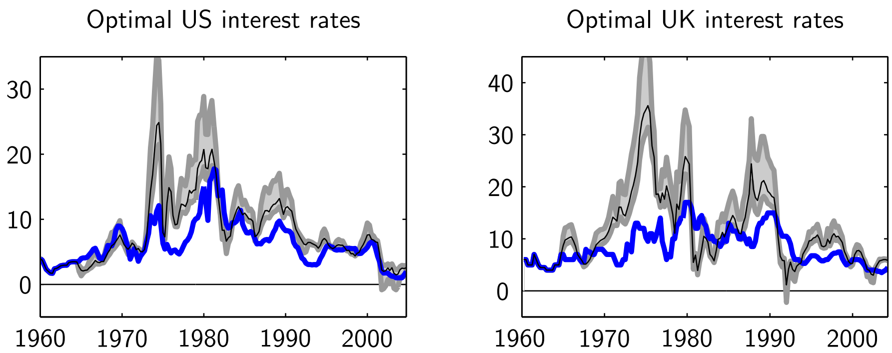

Figure 1.10: Optimal interest rates for varying weight on the interest rate smoothing component, $0.25 \leq \omega_\nu \leq 4$.

Policymakers with different weights on interest rate smoothing. For the second relative weight in the loss function we choose the same range, $\omega_\nu \in [1/4, 4]$. As in the previous exercise, the change in the relative weight on interest rate smoothing has a small effect on optimal rates, cf. figure 1.10. Policymakers judging variation in the interest rate very differently will nonetheless qualitatively agree on optimal rates, although of course policymakers with a low smoothing objective will allow for higher interest rates in the wake of the GI, in particular in the UK model where recommended rates reach very high levels in the mid–seventies consistent with the very high levels of inflation, while even a high smoothing objective has little effect on optimal rates given the high inflation rates.

Results for various degrees of inflationary bias. Figure 1.11 plots optimal policy decisions from policymakers that differ only in their tendency to allow for an inflation bias, that is in the choice of their preference parameter κ. We allow this parameter to vary over the whole admissible range from zero to one. The range of optimal interest rates widens substantially as we

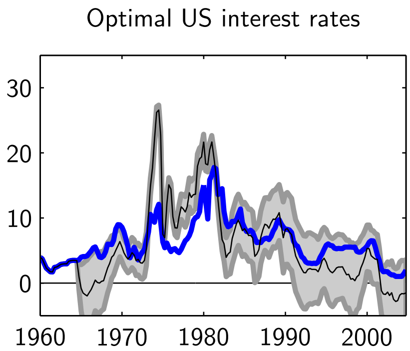

Figure 1.11: Optimal interest rates for different degrees of inflationary bias, $0 \leq \kappa \leq 1$.

include the paths pursued by fictitious central bankers with a strong tendency to reach unnaturally low unemployment rates at the cost of higher inflation. Interestingly, the period which is interesting for our question changes little: as inflation continues to rise, recommended rates approximately coincide with those from our benchmark simulation. Surprisingly, the inflation bias has little impact on optimal UK rates, as shown by the tight band of optimal interest rate paths.

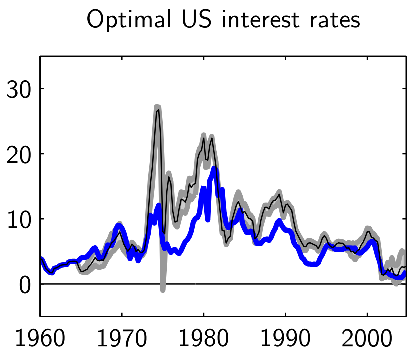

Figure 1.12: Optimal interest rates for different learning algorithms as specified by the gain parameter, $0 \leq \gamma \leq 0.06$.

Sensitivity towards the learning gain parameter. The gain value reflects two opposing forces: one is the desire to put higher weights on more recent observations to improve estimates in presence of structural breaks; in this regard higher gain values perform better, given that the learning dynamics are substantial as discussed earlier; the other is the desire to use sufficient sample information; if the gain value is too high there are effectively too few observations included in the estimates. The accord in the learning literature is on a gain value of $\gamma = 0.03$, but values around this particular value are equally reasonable. Specifically, we allow for gain values in the range $[0, 0.06]$. It should be noted, that with our non–recursive, discounted least squares formulation, a zero gain coincides with ordinary least squares, cf. appendix A.1, so our analysis nests all relevant learning specifications.

One might suspect that results are largely driven by the way learning is modelled, but as figure 1.12 confirms, quantitative differences in the estimates resulting from different gain parameters have little effect on optimal interest rate paths. This is an interesting observations since it documents that the assumption of adaptive learning per se has an important impact on our view of historical events, regardless of the particular specification of the learning algorithm.

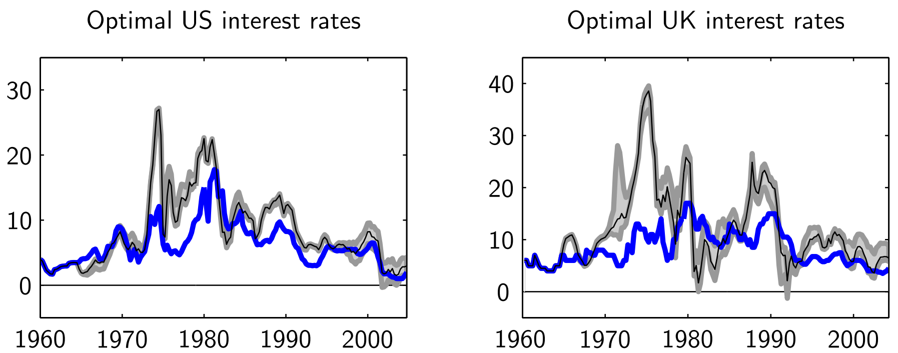

Figure 1.13: Optimal interest rates for different assumptions on the HP smoothing parameter used to extract the Nairu; among the values under consideration is the standard value 1600 but also the values that yield smooth Nairu paths as discussed in section 1.3.

Changing the perceived smoothness of the Nairu. We used a HP smoothing parameter that reproduced the time variation in the Nairu that is documented in the empirical literature, in particular Gordon (1997) for US data and Batini and Greenslade (2006) for UK data. However, we might also choose the more prominent value of 1600 for quarterly data or other lower as well as higher values than in our basic calibration. Specifically, we consider values 10^c with $c \in [3.2, 6.7]$. This exponential formulation allows us to put more weight on values that yield smooth series which we consider more plausible, as discussed in section 1.3, with the first entry giving the standard value.

Figure 1.13 plots the policy paths resulting from this exercise. As for the other parameters the HP smoothing value has little effect on optimal rates, the most notable differences appearing around the turning points of the Nairu, where differences in the smoothing parameter temporarily affects the perceived unemployment gap in different ways.

Apart from these few dates, the effect of decreasing μ is a more volatile perceived Nairu path in the first place. With this variable tracking the unemployment rate more closely, the resulting unemployment gap estimates become smaller in magnitude, but they remain qualitatively comparable to gap estimates resulting from lower values of μ. That is, all perceived paths of the unemployment paths obtained for different smoothing parameters are highly correlated and differ mainly in their amplitude. On the other hand, with the amplitude being low, the estimates of the Phillips curve slope will be high so that the product of both will be qualitatively the same across various smoothing parameters.

It should be noted, however, that a very volatile Nairu path obtained e.g. for the commonly used smoothing parameter $\mu = 1,600$ implies that the policymaker cannot avoid most of the variation in the unemployment rate, only the residuary and small deviations from it, and that even those departures of unemployment from the Nairu, misleadingly perceived as small, are thought to be highly effective. Thus, policymakers will recognize themselves being able to stabilize inflation with very little costs in real activity. This conjecture shall be analyzed in our companion paper where we verify these considerations.

Results under different assumptions on the policy impact coefficient. As argued in section 1.3, it is helpful for the analysis under adaptive learning to fix at least one of the parameters, the impact of the policy instrument on real activity, λ. However, this section is intended to demonstrate that this restriction does not significantly influence our findings. As argued above, some researchers use low values for this parameter, e.g. $\lambda = 0.024$ in Primiceri (2006), while other estimates indicate values of -0.035 (cf. Smets (2002)), and standard calibrations (Orphanides (2002)) suggest higher values around $\lambda = 0.08$. We will thus consider the whole range implied by these different values. Figure 1.14 plots the policy paths resulting from this exercise. As is obvious from this figure, our conclusions hold regardless of the specific value, though very low impact coefficients tend to invoke more aggressive policy decisions as their effect is limited by the low coefficient. The quantitative impact of changes in these parameters appears limited and optimal interest rate paths change little except for slightly higher rates at the peak dates.

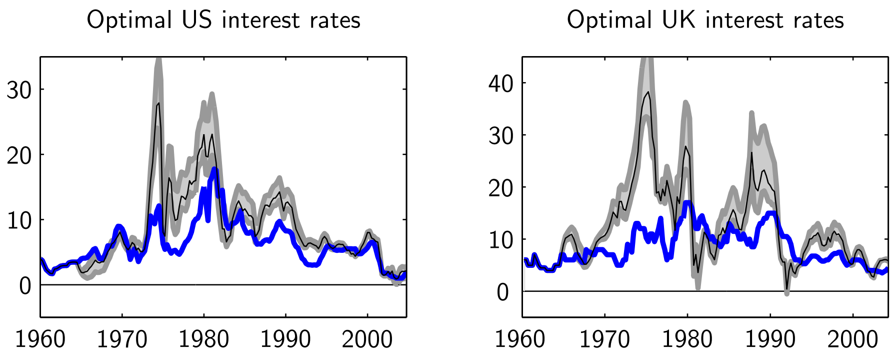

Figure 1.14: Optimal interest rates under different assumptions on the policy impact coefficient, $0.024 \leq \lambda \leq 0.08$.

Summary of the results. While the interest rate paths naturally differ for variations in the underlying key parameters, the finding that optimal rates exceed those historically observed holds for all specifications. Our robustness analysis demonstrates that it is the assumption of imperfect knowledge overcome by continuous data observation that yields our results, not any particular set of preference parameters or the specifics of the learning algorithms.

1.7. Conclusions

Our research question was to assess whether historical interest rates in the US and the UK are consistent with an optimizing but learning policymaker.

We found that there are no policy preferences that are entirely in line with those rates. In particular during the Great Inflation episode optimally recommended rates were substantially higher than those observed in the data, while in subsequent periods the interest rates set by the Fed and the Bank of England coincide with those of an optimizing learner. Nevertheless, for the whole sample we found a high positive correlation of US and UK interest rates with optimal rates. This high correlation is also present during the Great Inflation episode where actual rates comoved with optimally prescribed rates but were more conservative, while at the same time the standard deviation of the relevant estimates was substantially higher than in the following period. This suggests a connection with Brainard's (1967) conservatism principle that calls for such attenuate action in such a framework. Our results thus hint at the importance of integrating choice under uncertainty into the adaptive learning methodology.

Our robustness analysis demonstrated that the specification matters little for our results. Not only are the qualitative insights unchanged but even the quantitative impact of substantial changes in preferences and other key parameters such as the learning speed do not exert a substantial influence on our findings.

Our second result concerns the common view that policy considerably changed when Paul Volcker became chairman of the Fed. While we consider the same policymaker being in place for the whole period under consideration, his beliefs of key parameters and hence the response to target variables changes endogenously. This sheds new light onto the debate on whether the disinflation can be attributed to a change in the Fed's chairmanship, an argument that is for example advanced by Clarida et al. (2000). In our model every optimizing policymaker would have revised his stance on inflation and unemployment stabilization at around the time when Paul Volcker became chairman of the Fed, while conversely in the seventies any policymaker would have conducted policy in a way that would nowadays be criticized for having been non–optimal. This is an interesting result because it attributes the improvement in policy to endogenous forces rather then an exogenous change and thus applies equally well for different countries such as the UK which–as documented in the section presenting the learning results–faced qualitatively the same evolution of beliefs.

The fundamental question that still remains is whether the GI could have been avoided if the policymaker had acted more in line with optimally prescribed rates. This research question is taken up in the second chapter of this thesis, along with an investigation of the role Brainard–type uncertainty plays in our framework.

CHAPTER 2

Did the Great Inflation occur despite optimal policy?

2.1. Introduction

The term Great Inflation (GI) refers to the prolonged high inflation period experienced in the United States, and similarly in other industrialized countries, for almost twenty years. Beginning in the mid–sixties, US inflation gradually increased, peaking at double digit levels in the mid–seventies and remained high for a sustained period, until it finally returned to low levels under Federal Reserve chairman Paul Volcker. This rapid decline in the early eighties is now referred to as Volcker Disinflation. A large amount of research has been conducted to explain this episode, but until now it seems difficult to explain all relevant empirical facts consistently in a model. This introduction gives a brief description of these facts and reviews some representative explanations that have been advanced in the literature.

Beginning in the 1960s, US inflation gradually increased, reaching a maximum of 12 percent in 1974. Although it decreased for a while, it remained high on average and peaked again above 10 percent by the end of 1980. This second peak was followed by a sharp disinflation, which quickly brought the inflation rate below 4 percent within two years and to an average level of 2.5 percent thereafter. Unemployment lagged inflation, peaking one year and two years, respectively, after inflation.

One prominent view expresses doubt on whether policy was conducted properly at the time. Clarida et al. (2000) argue in a New Keynesian model that monetary policy during that episode can be described as following a Taylor rule of the form

$$i_t = i^* + g_\pi(E_t\pi_{t+1} - \pi^*) + g_x x_t, \tag{2.1}$$

It is well known in the literature, that rules of this type imply an indeterminate system for a large set of parameter constellations (g_π, g_x) in these models. By using wrong Taylor rule parameters and thus admitting self–fulfilling expectations, an inappropriate policy contributed to the Great Inflation.

Christiano and Gust (1999) express doubt on whether the stagflationary episode with jointly high inflation and unemployment can be explained in a New Keynesian model, as indeterminacy is associated with weak responsiveness to expected inflation, $0 < g_\pi < 1$. If expectations rise for some reason, the real interest rate falls, since the nominal rate rises by less than the increase in expectations. This lower real rate stimulates aggregate demand and, via the Calvo (1983) pricing assumption, produces upward pressure on prices, accommodating the initial increase in expected inflation and thus establishing them as self–fulfilling. Hence high inflation occurs in the New Keynesian Model simultaneously with low unemployment. The authors show, how in the Limited Participation model developed by Christiano et al. (1997), the self–fulfilling prophecy hypothesis can contribute to the debate. In this model, an initial rise in expectations and the associated fall in the real rate lead to reduced savings of the private sector. As a consequence less money is deposited in the financial market, in which firms have an unchanged demand for liquidity, which they need to finance their expenditures. This creates upward pressure on the nominal rate, which must be accommodated by the central bank by a monetary injection. In this model the rising nominal interest rate depresses investment expenditure and thus real activity, while the monetary expansion fuels inflation. This setup thus predicts the stagflation that is present in the data.

Orphanides and van Norden (2002) add an important observation to the discussion. He points out that policy makers base their decisions on real time data, while most of the literature, e.g. Clarida et al. (2000), uses revised data that became available several years after policy decisions were taken.[1] In this regard, one has to use real time data if one's goal is to analyze the appropriateness of policy decisions. Orphanides and van Norden suggest an intermediate concept, referred to as "quasi–real time data" that makes use

[1] The distinction between real time and quasi–real time data is important. Both concepts make use of samples that were available in the period in question, but while real time data also uses vintage time series, quasi–real time data considers segments of revised data series, as we do in this work since the latter are easily available while the former are usually not.

of revised data series but allows inference only from subsamples ending in the periods in which the corresponding decisions are made. In this spirit their analysis reveals that a determinate Taylor rule was in place, given the quasi–real time data, but that measurement problems, in particular related to the output gap, contributed to wrong policy decisions, and an a posteriori indeterminate system.

Goodfriend (1993) emphasizes credibility issues. A lack of commitment on part of the central bank led inflation expectations to become uncoupled from the underlying system, a phenomenon he refers to as inflation scares.

A common feature of the above interpretations is the central role of expectations, in particular those of inflation. Another common feature, however, is the impact of exogenous factors. Either the policymakers' preferences shift exogenously in advance of the Volcker Disinflation, as in Clarida et al. (2000) or expectations fluctuate exogenously, as in the explanation of Goodfriend (1993). Although it seems clear that expectations crucially depend on the economic environment and in particular on monetary policy, the above models are not capable of explaining why expectations did not stay well anchored. More recent research endogenizes the formation of expectations and is able to explain the evolution of beliefs, similar to the inflation scares as depicted by Goodfriend.

Orphanides and Williams (2005) present a model with private agents, who continuously update their beliefs concerning unknown parameters of the model and the unobservable Nairu, which is the rate of unemployment that induces no movement of inflation, frequently referred to as the non–accelerating inflation rate of unemployment, a term which led to the acronym Nairu. A key element in their analysis is the use of real time data. As the policy makers face measurement problems, misperceptions appear unavoidable. Expectation formation is assumed to be rational, apart from the need to make inference on unknown parameters. The learning mechanism, however, provides an additional propagation mechanism that causes parameter beliefs and hence expectations to fluctuate, thus endogenizing Goodfriend's inflation scares. They simulate inflation and unemployment paths under the hypothesis that either the Nairu was known or that the expectation formation mechanism remains fixed. As a result they find that absent either Nairu misperceptions or learning elements, the high and prolonged inflation would not have occurred.

In a recent paper Primiceri (2006) offers a coherent explanation of the Great Inflation in a New Keynesian model with an optimizing but learning policy maker. His model explains all empirical features, including the gradual initial rise in inflation, the sudden Volcker disinflation and the fact that inflation leads unemployment. A detailed description of Primiceri's model is offered in the next section. It is complementary to the work of Orphanides and Williams (2005, 2007) in that these authors demonstrate that the Great Inflation would not have occurred under perfect information, whereas Primiceri shows that it is the most likely consequence of imperfect knowledge.

2.2. Explaining the Great Inflation

Primiceri's model offers an economically plausible explanation for the Great Inflation. Starting point is the observation that in the 1960s the estimates of the Nairu were too low. Consequently, unemployment was mistakenly perceived to be significantly above satisfactory levels, prompting for stimulative monetary policy to promote economic recovery. At the same time, the inflation process was perceived as strongly mean–reverting, with a low degree of persistence. The rise in inflation was thus initially believed to be transitory and hence acceptable. The true inflation process, being highly persistent, caused an overly stimulative policy to push inflation further upward. Inflation peaked at a level of 12 percent in 1974:4 and remained high for a sustained period. Slowly, policy makers began to realize the persistent nature of inflation.

At this point estimates of the Nairu were still relatively low, so any attempt to control inflation via a positive unemployment gap appeared to fail: while the true gap was small and so was the implied reaction of inflation, the perceived gap was large, so the moderate reaction of inflation was attributed to a low impact coefficient. With an estimated Phillips curve slope near zero, it seemed virtually impossible (or unacceptable) to reduce inflation by admitting the required higher unemployment. Thus policy, assessing its situation as incapable of curing the economy, remained inactive.

On the other hand, with the high degree of persistence detected by now in the inflation process, even small changes in the perceived trade–off would enforce strong actions. This is what Primiceri suggests happened during the tenure of Paul Volcker. A sequence of favorable shocks revealed that the inflation–unemployment tradeoff was slightly higher than previously anticipated, inducing quickly rising interest rates. Hence, unemployment was pushed sufficiently above the true Nairu, causing the true Phillips curve mechanism to settle down inflation. As soon as this had been accomplished, the monetary authority allowed unemployment to return quickly to the neutral level given by the Nairu.

This interpretation can explain the slow and persistent rise in inflation (it was perceived transitory and thus not worth sacrificing unemployment, which was perceived to be high anyway), the prolonged period in which inflation was not brought back to lower levels (the policy maker felt incapable of reducing it at acceptable costs), and the fact that inflation leaded unemployment (unemployment was held high until inflation was sufficiently low).

2.3. Model, Estimation, and Learning

This section discusses our dynamic model of inflation and unemployment, and estimates variants of it after a brief digression on variable transformations used in the empirical analysis. Finally, we investigate the learning algorithm used by the policymaker.

2.3.1. The Model. We set up a reduced form of the New Keynesian model as discussed by Woodford (2003) where we identify real activity as the deviation of unemployment from the Nairu, $x_t \equiv u_t - u_t^*$, as in Primiceri (2006) and the previous chapter, where expectations have been replaced by backward looking terms,

$$\pi_t = c_\pi + \alpha(L)\pi_{t-1} + \theta(L)x_{t-1} + \varepsilon_t^\pi \tag{2.2a}$$

$$x_t = \rho(L)x_{t-1} + \lambda(r_{t-1} - \bar{r}) + \varepsilon_t^x \tag{2.2b}$$

While the theoretical model is specified with a lag length of one, empirical studies usually use longer lags structures in the pricing and demand equations, cf. e.g. Gordon (1997) for US data and Greenslade et al. (2003) as well as Batini and Greenslade (2006) for UK data, so we follow Primiceri in choosing an order of two for the lag polynomials, balancing both theoretical scarcity and empirical needs. Such a model is standard in the empirical literature and in the adaptive learning literature in particular, cf. Orphanides and Williams (2005) among others.

2.3.2. Estimation and Demeaning. Theoretical models of this sort are defined in terms of deviations of inflation and the policy instrument from their steady state values. As is also recognized by Primiceri (2006), the two constants and the level of the Nairu are not jointly identified. Empirical analysis usually resolves this issue by identifying equilibrium values by the sample means, referring to this as demeaning prior to estimation. Since we are considering a relatively long sample that covers the unusually high rates from the GI episodes, the sample means appear inappropriate as approximations for the equilibrium values. For a comparable sample, Smets (2002) uses an arbitrary linear detrending method for inflation while Primiceri treats the Nairu as stationary and fixes its equilibrium value at the sample mean of unemployment.

The standard demeaning procedure would implicitly subtract the sample means from inflation, which are given by 3.5 percent for US data and 6.4 percent for UK data. At least in a stationary environment this would imply that without policy interference inflation would stabilize around these high values eventually. Any attempt to attain a lower equilibrium rate would necessitate to permanently hold unemployment above the Nairu to assure a balance between his inflation and unemployment goals. It appears more reasonable to assume that the equilibrium inflation rate is consistent with the policymaker's target rate, hence we impose these values instead of relying on sample means. The same applies for the real rate. Therefore, we calibrate the equilibrium value of the real rate of interest at $\bar{r} = 3$ percent and identify the equilibrium inflation level with the policymaker's target value for inflation, $\pi^* = 2$ percent, which appears consistent with other empirical studies, cf. e.g. Orphanides and Williams (2005) and Gerlach and Svensson (2003), and actual central bank behavior. This expresses the presumption that the goal set by the policymaker can be achieved in a dynamically consistent way.

2.3.3. Estimation Results. Before we can simulate paths for inflation and unemployment, we need to estimate the model which we will use as the true data generating process. We will do so by estimating an appropriate state space form of our basic equations, depending on the specification we wish to use. We shall use two different scenarios, one being the unrestricted model and the other featuring a unit root in inflation.

To jointly estimate the path of the Nairu and the parameters of the model, we need to set up an appropriate state space form.

Collecting the observable variables in the measurement vector $m_t = (\pi_t, u_t)'$, and the unobservable variables in the state vector $s_t = (u^*_t, u^*_{t-1}, u^*_{t-2}, \alpha_{1,t}, \alpha_{2,t})'$, we can estimate our benchmark model by maximizing the likelihood of the state space model

$$m_t = \begin{pmatrix} 0 & -\theta_1 & -\theta_2 & \pi_{t-1} & \pi_{t-2} \\ 1 & -\rho_1 & -\rho_2 & 0 & 0 \end{pmatrix} s_t + \begin{pmatrix} \theta_1 & \theta_2 & 0 \\ \rho_1 & \rho_2 & \lambda \end{pmatrix} x_t + \begin{pmatrix} \varepsilon^\pi_t \\ \varepsilon^x_t \end{pmatrix} \tag{2.3}$$

$$s_t = \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 1 & 0 & 0 & 0_{3\times2} \\ 0 & 1 & 0 & \\ \hline \multicolumn{3}{c|}{0_{2\times3}} & I_2 \end{array}\right) s_{t-1} + \left(\begin{array}{c} \varepsilon^*_t \\ 0 \\ 0 \\ \hline 0 \end{array}\right) \tag{2.4}$$

with the exogenous vector $x_t = (u_{t-1}, u_{t-2}, r_{t-1})'$ and $r_t \equiv i_t - \bar{\pi}_t$ where $\bar{\pi}$ is a four–quarter moving average of current and past inflation as in Rudebusch and Svensson (1998). The unit root version incorporates the restriction $\alpha_1 + \alpha_2 = 1$, cf. Hamilton (1994) for a discussion, so we estimate the modified state space model, now in the state vector $s_t = (u^*_t, u^*_{t-1}, u^*_{t-2}, \alpha_{1,t})'$

$$m_t = \begin{pmatrix} 0 & -\theta_1 & -\theta_2 & \Delta\pi_{t-1} \\ 1 & -\rho_1 & -\rho_2 & 0 \end{pmatrix} s_t + \begin{pmatrix} \theta_1 & \theta_2 & 0 & 1 \\ \rho_1 & \rho_2 & \lambda & 0 \end{pmatrix} x_t + \begin{pmatrix} \varepsilon^\pi_t \\ \varepsilon^x_t \end{pmatrix} \tag{2.5}$$

$$s_t = \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 1 & 0 & 0 & 0_{3\times1} \\ 0 & 1 & 0 & \\ \hline \multicolumn{3}{c|}{0_{1\times3}} & 1 \end{array}\right) s_{t-1} + \left(\begin{array}{c} \varepsilon^*_t \\ 0 \\ 0 \\ \hline 0 \end{array}\right) \tag{2.6}$$

with the exogenous vector $x_t = (u_{t-1}, u_{t-2}, r_{t-1}, \pi_{t-2})'$.

Table 2.1 supplies the estimation results for the two versions for US and UK data, and figure 2.1 plots the corresponding estimates of the Nairu paths.

	US model		UK model	
	unrestricted	unit root	unrestricted	unit root
α_1	0.646 [8.52]	0.672 [9.00]	1.384 [22.29]	1.414 [22.54]
α_2	0.298 [3.94]	0.328 [—]	-0.427 [-6.88]	-0.414 [—]
θ_1	-0.935 [-2.65]	-1.081 [-3.04]	-1.060 [-3.12]	-1.353 [-3.88]
θ_2	0.811 [2.29]	0.983 [2.76]	0.950 [2.69]	1.318 [3.66]
ρ_1	1.667 [25.08]	1.670 [25.17]	1.939 [26.09]	1.940 [25.72]
ρ_2	-0.711 [-10.78]	-0.715 [-10.89]	-0.999 [-13.59]	-0.999 [-13.38]
σ_π^2	1.379 [8.65]	1.394 [8.55]	1.150 [9.01]	1.168 [8.80]
σ_x^2	0.045 [6.02]	0.045 [6.02]	0.050 [5.54]	0.051 [5.42]
$\mathcal{L}$	-270.26	-271.56	-325.84	-328.47

Table 2.1: Estimation results for US and UK data for the unrestricted and unit root versions. T–statistics in brackets.

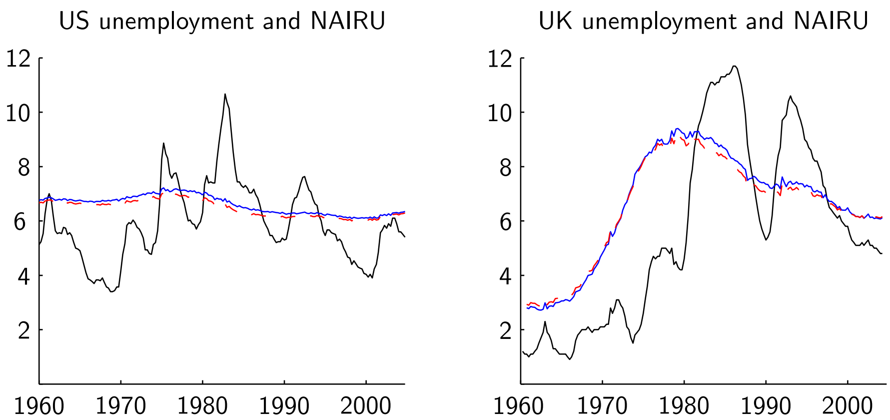

Figure 2.1: Estimated Nairu paths in the unrestricted model (solid lines) and the unit root model (dashed lines).

2.3.4. Learning. This section intends to give a brief review of the mechanism of learning. For a more detailed exposition, the reader is referred to the corresponding section of chapter 1. As discussed there, the policymaker is assumed to revise his perceived law of motion,

(2.7a) $$\pi_t = \hat{c}_\pi + \hat{\alpha}_1 \pi_{t-1} + \hat{\alpha}_2 \pi_{t-2} + \hat{\theta}_1 x_{t|t-1} + \hat{\theta}_2 x_{t|t-2} + \varepsilon_t^\pi$$

(2.7b) $$x_t = \hat{c}_x + \hat{\rho}_1 x_{t|t-1} + \hat{\rho}_2 x_{t|t-2} + \lambda(i_{t-1} - \bar{\pi}_{t-1}) + \varepsilon_t^x,$$

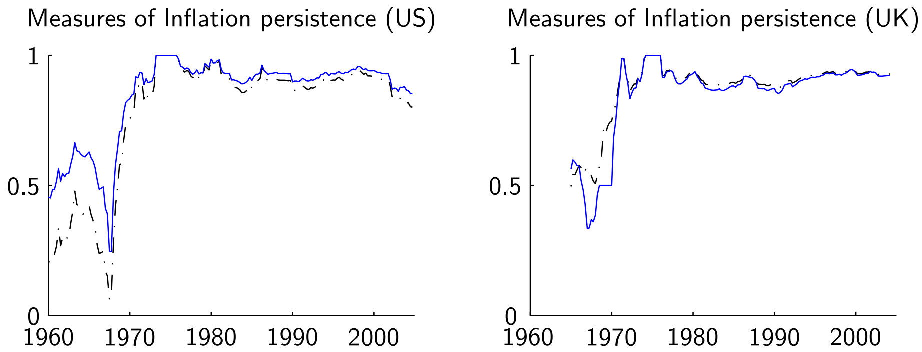

Figure 2.2: Perceptions of inflation persistence as measured by the largest eigenvalue (solid lines) and the sum of AR(2) coefficients (dashed lines).

in each period. Firstly, he extracts the Nairu with the HP Filter. After computing the unemployment gap as difference between unemployment and the Nairu, he may choose to discount all regressors and endogenous variables with a geometrically declining scaling factor, or leave them in their original form. Depending on this choice he computes the least squares estimates in either the transformed or the original model, the former method being referred to as discounted least squares, cf. Harvey (1993). The choice of a discount factor δ which is used to scale an observation s periods in the past by δ^s, corresponds to a method frequently applied in the adaptive learning literature to reflect agents' concern about structural changes. The related learning algorithms are referred to as constant gain algorithms, and they are recursive formulations of discounted least squares. Chapter 1 and appendix A.1 provide a more detailed account of this topic.

Over time, the central banker accumulates knowledge on the Nairu and the model's parameters. As his information set grows, his perceptions about economic relations evolve over time, and in each period imply an updated transition law of the economy that he uses to set interest rates optimally. Before proceeding to a brief description of his decision problem, we examine a few summary statistics regarding his beliefs.

Of particular interest is the change in policymakers' perceptions regarding inflation persistence and the path of the Nairu over time. Figure 2.2 depicts the estimates of the sum of autoregressive coefficients in the Phillips curve, $\alpha_1+\alpha_2$ (dashed line). This sum is a standard reference value to assess persistence

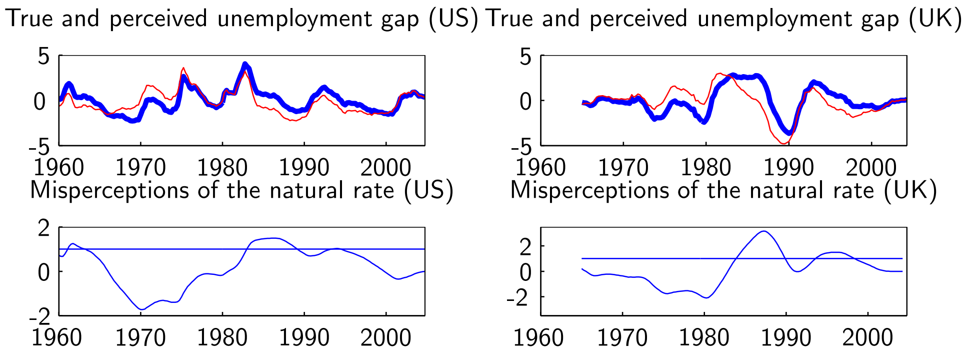

Figure 2.3: Upper panels depict the true (thick line) and perceived unemployment gaps, the bottom panels illustrate the Nairu misperceptions.

in the inflation rate. As can be seen, inflation was seen as a strongly mean–reverting process until the mid–seventies while thereafter agents became aware of its highly persistent nature, although it is viewed as stationary in most subsamples. It can be argued that a more appropriate measure of persistence is the dominant eigenvalue of the Phillips curve. Figure 2.2 also contains the evolution of this measure (thick line). Apparently, both measures give a very similar idea of inflation persistence.

The Nairu estimates change substantially over time, as is equally documented by Orphanides and van Norden (2002) and our previous analysis in chapter 1. Early quarters in our sample are characterized by a large and persistent underestimation of the Nairu, which for the prevailing unemployment rates pushes real activity more in the focus than it should have in retrospect. For a graphical illustration, see figure 2.3. This figure plots the US and UK unemployment gaps as perceived in real time (dashed) and based on the likelihood estimates of the Nairu, which we refer to as true Nairu.

Interestingly, the Nairu misperceptions in the US and the UK model share strong qualitative similarities, strongly comoving and being of the same sign that they switch at around the same dates, though of course their amplitude differs owing to the different smoothness assumptions imposed on the Nairu path.

Orphanides and Williams (2005) construct an analogous series of US Nairu misperceptions based on narrative evidence, and despite the use of a different methodology present remarkably similar results. They also report initial negative misperceptions, which imply an overly optimistic view about low levels of the Nairu, which were falsified in retrospect. They support our finding that the level of the Nairu was persistently underestimated until the early eighties.

The policymaker sets nominal interest rates to minimize a standard infinite sequence of quadratic period losses,

$$\sum_{s\geq t}\beta^{s-t}[(\pi_s-\pi^*)^2+\omega_u(u_s-k\hat{u}^*_{t|s})^2+\omega_\nu(v_s-v_{s-1})^2] \tag{2.8}$$

which punish deviations of inflation from the target value π^*, changes in the interest rate, and deviations of unemployment from a target level that equals a fraction κ of the Nairu. The case in which this fraction is one corresponds to the standard case, while all other parameter choices for κ imply the desire to push unemployment below the Nairu, as in the Barro and Gordon (1983) model, where surprise inflation is used to attain higher than normal levels of real activity. This parameter gives us the flexibility to account for suboptimal policy behavior associated with the inflation bias associated in Barro and Gordon's model.

Given that the central bank's model of the economy changes each period, it will have to recompute the optimal feedback to the state of the economy in each period. For fixed policy preferences and data series that we sequentially generate with the model equations and the estimates from table 2.1, as described in more detail below.

2.4. Simulation Results

This section presents our simulation results of an adaptively learning but optimizing policymaker. We precede the main part which presents the simulation in detail by a discussion on the benchmark without learning. The following section will then conduct robustness checks on our basic learning exercise and a final part considers the effect of incorporating findings from chapter 1 into the analysis and find that a combination of adaptive learning and policy conservatism replicates the prolonged and very high inflationary episodes in the United States and the United Kingdom.

2.4.1. Benchmark scenario without learning. To contrast our adaptive learning results with the no learning case we consider the case of perfect information, where the policymaker has knowledge of all relevant variables and parameter values. The true data generating process can be parameterized by a vector $\xi^* = (\alpha_1^*, \alpha_2^*, \theta_1^*, \theta_2^*, \rho_1^*, \rho_2^*)'$ containing the unrestricted estimates of the models from table 2.1. We consider the case of no inflationary bias, i.e. $\kappa = 1$, since this yields a stationary model, with optimal policy feedback parameters on the unemployment and Nairu terms summing to zero, so that we can reformulate the dynamics under optimal policy of a perfectly informed policymaker in form of a stationary autoregressive process as[2]

$$\tilde{y}_t = c + M\tilde{y}_{t-1} + \tilde{\varepsilon}_t \tag{2.9}$$

in the stationary variables $\tilde{y}_t = (\pi_t, \pi_{t-1}, \pi_{t-1}, \pi_{t-1}, u_t - u_t^*, u_{t-1} - u_{t-1}^*, v_t)$ and deduce mean and variance of inflation from the vector of unconditional expectations $E(\tilde{y}) = (I - M)^{-1}c$ and the unconditional covariance matrix $\tilde{\Sigma}$ which is implicitly defined by the Lyapunov equation $\tilde{\Sigma} = M\tilde{\Sigma}M' + \text{cov}(\tilde{\varepsilon})$.

Even under perfect information inflation can reach arbitrarily high values although higher values become less likely. Accordingly, we shall refer to a path of inflation exhibiting a Great Inflation only in case it peaks at unusually high levels, in the sense of exceeding the critical value, π_{crit}, that would be surpassed only with a small probability, say five percent, in our perfect knowledge benchmark. Formally, this value is defined as $\text{Prob}(\pi > \pi_{crit}) = 0.05$ with the probability measure being defined by the law of motion (2.9). Table 2.2 summarizes the relevant values for the US and UK benchmark models. Interestingly, historical inflation reached substantially higher values than would be expected in a no–learning environment.

2.4.2. Simulation of Optimal Dynamics under Adaptive Learning. We use the the first five years as training samples to determine initial beliefs, that is 1954:3–1959:4 for the US model and 1960:2–1965:3 for the UK model. For any subsequent date t the policymakers reestimate their perceived law of motion (2.7) for the given subsample and set up an optimal path for the

[2] In the case allowing for an inflationary bias, $k < 1$, the optimal policy feedback would also include a reaction to the level of unemployment and hence would inherit the non–stationarity of the Nairu. Primiceri offers evidence that the inflation bias term is not very important.

Model	Mean	Variance	Critical Value	Inflation peak
US model	2%	3.1%	7.0%	11.7%
UK model	2%	4.6%	9.5%	21.6%

Table 2.2: Unconditional mean, variance and a critical value of inflation–as defined in the text–under optimal policy and perfect information, as well as the historical peak of inflation during the Great Inflation.

current and expected future interest rates. After the current period rate is implemented, the underlying model equations (2.2) generate realizations of the endogenous variables for the next period. The policymakers observe these values and the process is repeated. As data generating process we will use both the unrestricted and the unit root version of the model.

In this section we will focus on benchmark values for the policymakers' preferences. In particular, we set $\pi_0^* = 2$ as in the estimation of the previous section, and choose a smoothing parameter for HP Filter that yields full sample Nairu paths comparable to our estimated paths from the previous section. As benchmark weights in the loss function we choose $\omega_u = \omega_\nu = 1$ so that all target variables are equally important. As benchmark value for the inflation bias parameter we set $\kappa_0 = 0.87$, the estimate Primiceri reports in his analysis. This value is already very high and precludes much of the motive to surprise inflate to generate particularly low levels of unemployment. The learning gain parameter is set to the standard value of $\gamma_0 = 0.03$, cf. Milani (2005) and Primiceri (2006). We will investigate changes to each of these parameters in the next section.

We simulate $n = 10,000$ series of inflation and unemployment and investigate whether they exhibit the high inflation episodes encountered in the US and the UK during the 1970s.

For convenience, we will denote the quarter in which the inflation reaches its maximum by t^*. From our discussion of the stylized facts during the great inflation, we are particularly interested in the highest levels the inflation rate attains during the simulation, and in the highest value of the unemployment gap surrounding t^*. Accordingly we gather the highest gap value that occurred in the period three years before or three years after t^*, and examine how many periods pass until the gap reaches its maximum, a positive value reflecting the situation encountered in the data where unemployment lags inflation. We will investigate the length of the disinflation period, which we define as the time that passes until inflation returns from its peak to its target value. A final focus is on the beliefs around t^* and the preceding periods. As discussed in section 2.2, the estimates of inflation persistence are low in the 1960s and early 1970s, while reaching unity at around t^*. We thus compare the highest estimates of persistence in the year ending in t^* with the mean value over the four preceding years. Similarly, we will analyze the slope of the Phillips curve. Since it fluctuates stronger and the occurrence of its maximum is more dispersed than the case of the persistence parameter, we measure the mean over the last 5 years before t^* and contrast it to the highest value the perceived slope attains in the three foregoing years.

In section 2.4.1 we argued that without learning the optimal dynamics can be expressed as the reduced form 2.9, at least in the case without inflation bias, i.e. $k = 1$. From this equation we deduced a critical value of inflation that will be exceeded in simulations only with a small probability, which we fixed on 5 percent. Under learning, the simulated paths of inflation will potentially exceed this critical values with a higher frequency, depending on the dynamics of the model under learning and the preferences of the policymaker.

We refer to the percentage of simulations with an inflation peak exceeding our critical value from the no–learning benchmark as the probability of a Great Inflation. It should be noted that this implicitly sets the probability of a Great Inflation in the no–learning scenario at five percent. It might be argued that this is a biased measure since we explicitly allow for an inflation bias, but our benchmark learning calibration of the relevant parameter is close to one ($\kappa_0 = 0.87$) and as we shall see in the next section, our result is only marginally sensitive to changes in this parameter. Figure 2.4 shows the distribution of the inflation peaks. The inflation peaks are high, but with median values between 8.5 and 9 percent in the US model and 11.7 and 12.6 in the UK model fifty percent of the simulations fall below the inflation rates actually experienced in the data, which were 11.7 in 1980:1 and 21.6 percent in 1980:3, in the US and the UK respectively. Only a quarter of the simulations surpass values of 9.9 and 14.2 percent, hence while the simulations of the US model coming close to

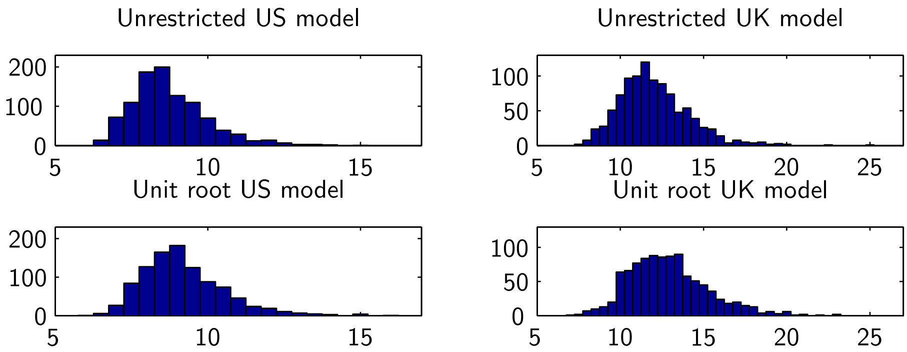

Figure 2.4: Histogram of inflation peaks in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

observed inflation peaks, the UK model is far from explaining the size of the great inflation in the UK, even if the outcomes are consistently higher than in the US model. We will discuss this issue in our final section that accounts for conservative policy behavior in the sense of Blinder (1997).

The probability of a GI is 93.8% percent in the unrestricted US model, 94.5% percent in the unit root US model, and 92.3% and 95.6% in the corresponding UK models. It is not surprising that in the unit root model the simulations result in more pronounced inflation outbreaks since the policymaker not only has incomplete information about the stochastic processes, he also ignores the non–stationarity in his considerations. One might conclude that the adaptive learning dynamics substantially contribute to explaining the high levels of inflation, but we are also interested in whether the characteristic features that jointly occurred in the actual economies are also present in our simulations.

One important aspect of the GI and the associated disinflation was the fact that unemployment was brought to high levels relative to the Nairu, the highest unemployment gaps having been observed at 3.2 and 3.0 percent in the US and UK time series, 11 and 6 quarters after the inflation peaks (t^*). This was associated as deliberate policy move to fight the high inflation by slowing down real activity until inflation was back to acceptable levels. We are thus also interested in the distribution of the unemployment gap around the

	Quartiles			Statistic	Quartiles		
Statistic	1^{st}	2^{nd}	3^{rd}	Statistic	1^{st}	2^{nd}	3^{rd}
Inflation Peaks	7.9	8.5	9.4	Persistence_0	0.45	0.72	0.85
	8.3	9.0	9.9		0.48	0.77	0.90
	10.6	11.7	13.1		0.71	0.84	0.90
	11.2	12.6	14.2		0.76	0.88	0.94
Gap Peaks	1.9	2.7	3.6	Persistence_1	0.92	0.99	1.00
	2.5	3.4	4.4		0.97	1.00	1.00
	1.9	2.5	3.2		0.93	0.98	1.00
	2.0	2.7	3.5		0.97	1.00	1.00
Quarters	4.0	5.0	8.0	Slope_0	−0.52	−0.36	−0.20
	4.0	5.0	7.0		−0.50	−0.33	−0.19
	2.0	3.0	4.0		−0.91	−0.62	−0.45
	2.0	3.0	4.0		−0.87	−0.60	−0.46
Disinflation	8.0	13.0	20.0	Slope_1	−0.32	−0.16	−0.03
	9.0	15.0	22.0		−0.29	−0.14	−0.03
	5.0	8.0	13.0		−0.55	−0.38	−0.14
	5.0	9.0	16.0		−0.54	−0.38	−0.18

Table 2.3: Summary of simulation results. The table supplies the quartiles for the variables of interest, where for each the top two lines correspond to the unrestricted and unit root US model, the bottom two lines contain the corresponding results from the UK models

inflation peaks. Figure 2.5 offers the corresponding histogram of the size of the gap, and figure 2.6 depicts the lags between inflation and gap peaks, a positive value indicating that inflation leads unemployment gaps. The simulations also yield perceived unemployment gaps peaking after inflation, in line with the idea that unemployment was raised above the Nairu level in order to bring inflation back to moderate levels. Table 2.3 provides the relevant statistics in the rows labelled 'gap peaks' and 'quarters'. The median values attained range from 2.7 to 3.4 percent in the US simulation covering the historical peak value, and they occurred 5 quarters after inflation reached its maximum, thus faster than in the US time series; however, values of 7 and 8 quarters are surpassed

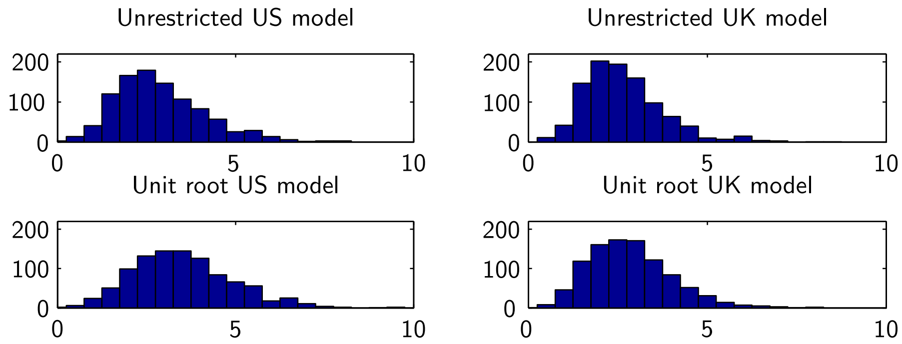

Figure 2.5: Histogram of unemployment gap peaks in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

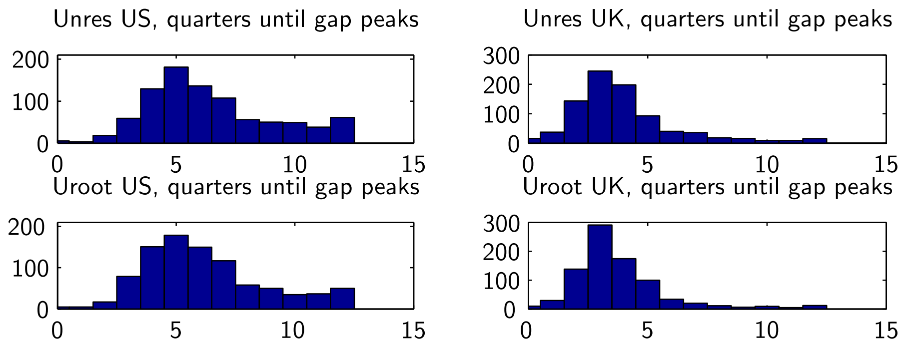

Figure 2.6: Histogram of period between inflation and unemployment peaks in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

by a quarter of the unrestricted and unit root model simulations. The UK simulations reached the maximum values of the perceived unemployment gap substantially earlier then the US counterpart, as we also find in the data. Fifty percent of the simulations peak before or in the third quarter after the inflation peak, 75 percent before or in the fourth quarter thereafter, which is slightly faster than observed in the data.

Another aspect was the rapid disinflation. Inflation continually rose for almost a decade, but after policy used its instrument to sharply control unemployment rates, it took only a few quarters until the disinflation successfully ended. We analyze how many quarters where needed until inflation reached

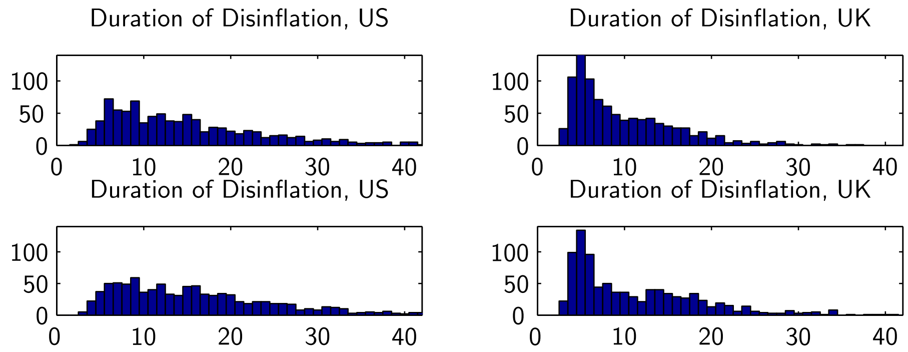

Figure 2.7: Histogram of duration of the disinflation period in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

its target value, and summarize the results in figure 2.7. US inflation reached moderate levels (below 3 percent) by the end of 1983, that is 12 quarters after the t^*, although it took until 1986 until inflation was below 2 percent for the first time. Our simulations confirm this finding in our US model under learning, since the disinflation took about 13 to 15 quarters, and even values of 20 to 22 quarters frequently occurred. UK inflation very quickly came down to moderate levels about two years after t^* and remained there thereafter, although it never reached the assumed 2 percent target in the 1980s. The model simulations yield a similarly rapid disinflation, occurring 8 to 9 periods after t^*, with a quarter of the simulations exceeding disinflation periods of 13 and 16 quarters length. We thus find that our simulations replicate the size and timing of historically observed unemployment gaps.

We argued that a crucial determinant of the great inflation outbreak was the evolution of beliefs about inflation persistence. It was low during the periods before the rise of inflation and with inflation approaching its maximum became closer to unity. The left panels of figure 2.8 plot the mean value of persistence over the four years preceding t^* and the right panels plot the highest estimates in the year ending in t^* for the US and UK simulations. The respective top panel corresponds to the unrestricted model, the bottom panel to the unit root version. Note that the ordinate is cropped in the right panels for expositional purposes: the majority of persistence estimates at the heights of the GI equals unity, though some values are still slightly below. Since the

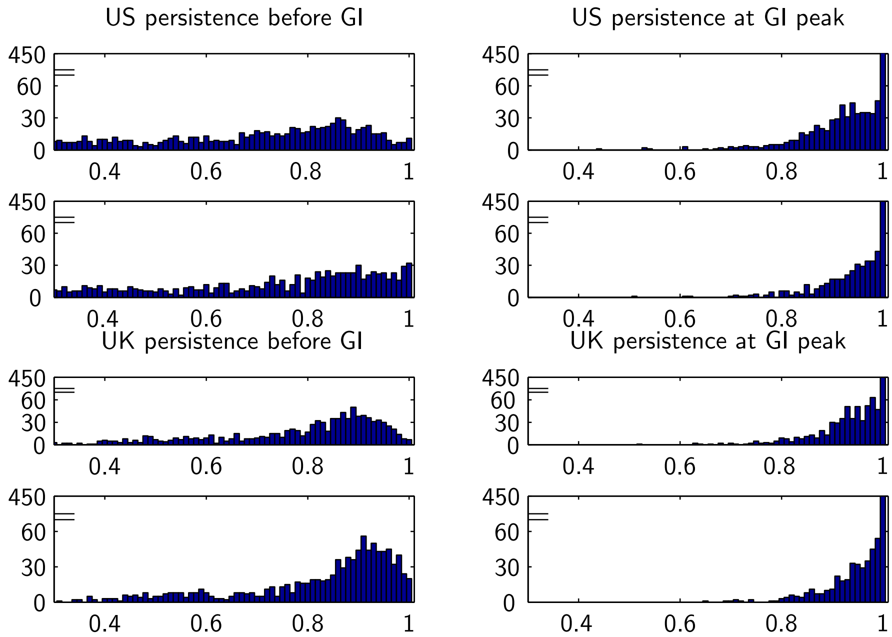

Figure 2.8: Histogram of perceived inflation persistence before and during the great inflation in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

histogram value at unity is ten times as large as at the remaining values, the ordinate values between 75 and 425 are cut, the only bar exceeding any of these two values being the outer right one at unity. We thus see that while the average estimates of inflation persistence are dispersed around low values (left panels) this perception sharply changes at the time of the inflation peak where most estimates indicate the near unit root property of the inflation series. Thus, as we already discussed for US and UK data, our simulations show a sharp and sudden increase of perceived persistence on the outbreak of the great inflation. Table 2.3 summarized this finding: the entries in the rows 'persistence$_0$' show that fifty percent of the simulations generated beliefs about this parameter in the ranges 0.45–0.85 (unrestricted US model) and 0.48–0.90 (unit root US model), as well as 0.71–0.90 (unrestricted UK model) and 0.76–0.94 (unit root UK model), whereas 75 percent exceeded values around 0.92 (unrestricted models) and 0.97 (unit root model), and 50 percent were almost unity, as summarized by the rows labelled 'persistence$_1$'.

As inflation rose the estimates of the Phillips curve slope were shown to be adjusted to nearly ineffective values. The same phenomenon also occurs in our simulations, though less pronounced: while the perceived slope has substantial (negative) values in the years preceding t^*, most estimates drop to almost zero when inflation reaches its peak. Figure 2.9 shows the distributions of perceived slope parameters, again the respective top panel depicts results from the unrestricted model, and the bottom panel refers to to unit root version of our model. The left panel contains histograms of the slope prior to the great inflation in period t^*, the right panels show how the perception changed in the periods around t^*. Evidently, the slope perceptions were distributed around relatively high negative values, but as inflation continuously rose until reaching its peak at t^*, policymakers became more and more sceptical about their ability to fight inflation via their impact on real activity. A substantial number of simulated slope estimates approaches zero, thus negating any influence of changes in the unemployment gap on inflation. Nevertheless, many estimates remain different from zero, a fact that is also reflected in the summary statistics contained in the rows 'slope_0' and 'slope_1' of table 2.3, which support the finding that the estimates became less favorable, but also document that the many estimates still remained non–zero.

2.5. Sensitivity Analysis

Our previous analysis focussed on a single type of policymaker characterized by our benchmark calibration. We will now investigate to what extend our results generalize in face of arbitrary policy preferences or changes in the modelling of the learning mechanism. Instead of repeating the full analysis of the preceding section we will concentrate on summary statistics in dependence of key variables, in particular we will pay attention to the quartiles of the relevant magnitudes, as listed in table 2.3 in the benchmark exercise. Besides the three policy parameters we shall also explore the role of the gain parameter in the learning algorithm and the parameter of the Hodrick–Prescott filter applied by the policymaker to extract the Nairu.

We analyze how these changes in preference or learning parameters changes our key results, the inflation peak, the highest unemployment gap attained to convey a disinflation, the length of the disinflation period, the number of

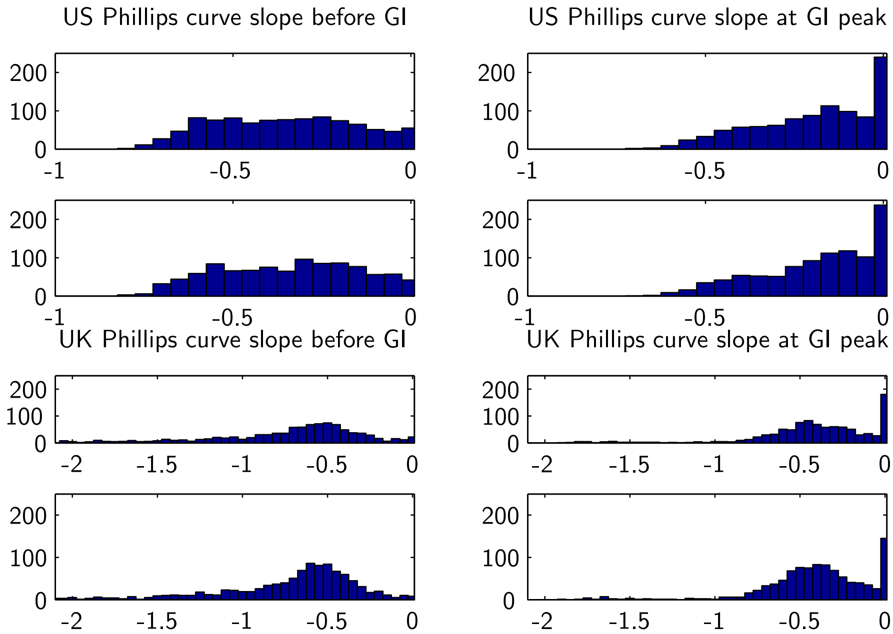

Figure 2.9: Histogram of perceived Phillips curve slope before and during the great inflation in the US models (left panels) and UK models (right panels). The top panels show results from the unrestricted model, the bottom panels those from the unit root versions.

period the gap lags or leads the inflation peak, and the probability that a Great Inflation period occurs given policymakers face imperfect knowledge.

It might well be the case that policymakers with different preferences or learning algorithms would not be confronted with GI type episodes, so that the historical high–inflation periods could have been avoided by resorting to a more appropriate policy stance. However, as it turns out, all optimizing policymakers would have found themselves trapped in a GI had they faced imperfect knowledge, at least with a very high probability. Thus in addition to the finding of the previous section that demonstrated that a GI is likely to arise under the requirement to learn about the economic environment, the results of this section support the notion that imperfect knowledge as such is the source of these disastrous economic outcomes, and not any particular specification of preferences.

Considering different weights on unemployment stabilization. Figure 2.10 as well as the corresponding ones in the following sensitivity exercises

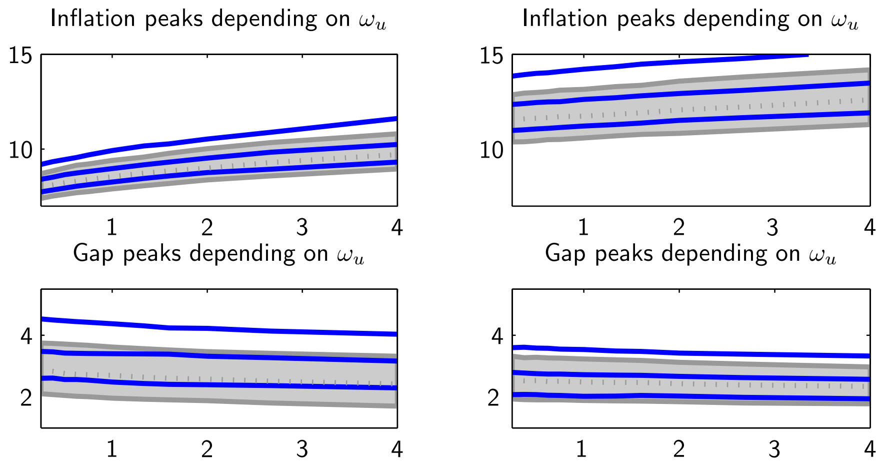

Figure 2.10: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different weights on unemployment stabilization

contain the first, second and third quartiles of inflation and the unemployment gap depending on the parameter in question–here the weight on unemployment stabilization, for both the unrestricted model, which are depicted as thick lines, as well as those for the unit root model, which are depicted as shaded regions.

The resulting maximum unemployment gap under optimal policy is little affected by a change in the relative weight on unemployment stabilization, although the curves have negative slope, as expected: the higher the weight on unemployment stabilization relative to inflation the lower is the resulting maximum deviation in unemployment from the (perceived) Nairu. Reconsidering the inflation peaks, this may seem puzzling, since the differences in the resulting inflation peaks are more pronounced. This can be explained in light

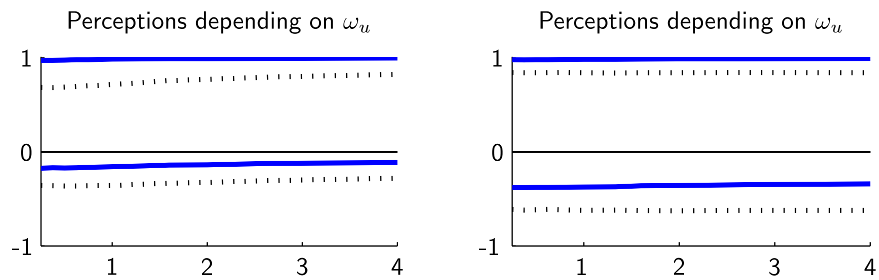

Figure 2.11: Policymakers' perceptions of inflation persistence and Phillips curve slope before and at the height of the GI for different weights on unemployment stabilization

of the evolution of the perceptions of the Phillips curve slope, $\theta(1)$: inspection of figure 2.11 reveals that close to the Great Inflation we observe a drop in the (absolute value of) the slope, that is, as in the benchmark case, policymakers substantially revise their slope estimate only a few quarters before the peak of the GI, and with higher weights on unemployment stabilization this estimate becomes more and more pessimistic. The reason for this might be the interaction with the Nairu misperception: the more reluctant the policymaker is to inducing movements in the unemployment rate the more likely it becomes that deviations in real activity are soaked up by movements in the underlying Nairu although – crucially – the policymaker is not aware of this due to his poor Nairu estimates; hence, a policymaker that operates with relatively low levels of (perceived) unemployment gaps is more likely to find his actions phasing out without any substantial effect on inflation, hence forcing him to revise his slope estimate towards less effective values. The top panel of figure

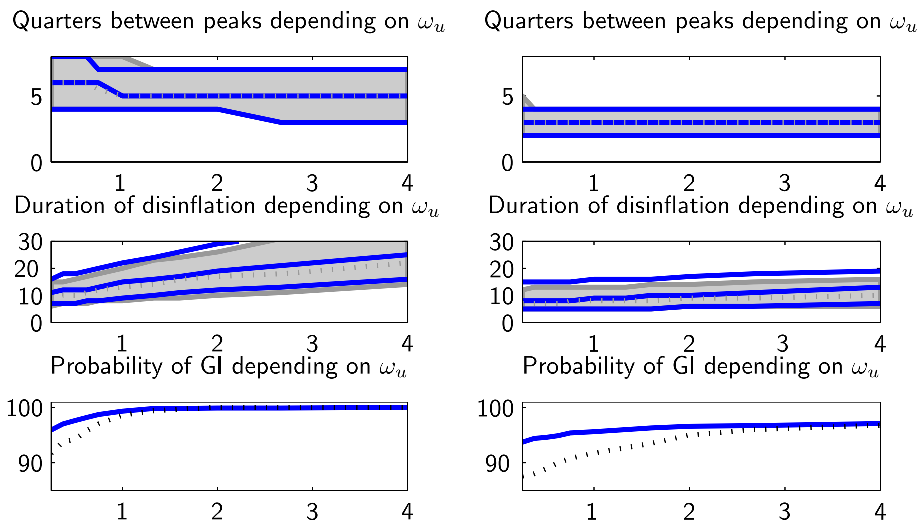

Figure 2.12: Periods that unemployment peak lags inflation peaks (top panel), duration of disinflation and probability of a GI for different weights on unemployment stabilization

2.12 summarizes the number of periods that the peak of the unemployment gap lags the peak of inflation, with negative numbers representing a lead of unemployment gap peaks. Again, the thick lines correspond to the results from the unrestricted model, the shaded regions to those in the unit root model. The results clearly indicate that the unemployment gap lags inflation with a five period delay in the US and with a three period delay in the UK, regardless of the choice of the weighting parameter. Similarly, the panels in the middle reflect the number of periods that are needed to bring inflation back to its target value. In the US, the duration of this disinflation period increases from a median value of 10 quarters to 20 quarters as unemployment stabilization becomes relatively more important. In the UK, this increase is less pronounced. Starting from 10 quarters for a low weight on unemployment stabilization of $\omega_u = 0.25$, the disinflation takes 12 quarters for a high weight of $\omega_u = 4$. The bottom panels depict our measure of the probability of a Great Inflation. As might be expected, the higher the relative weight the policymaker attaches to unemployment, the higher is the risk of high inflation periods. However, the occurrence of a Great Inflation is very likely, given the high values above 90 percent and in many cases even 100 percent. The quartiles are very close to each other and would be difficult to distinguish visually and therefore we only present the median values.

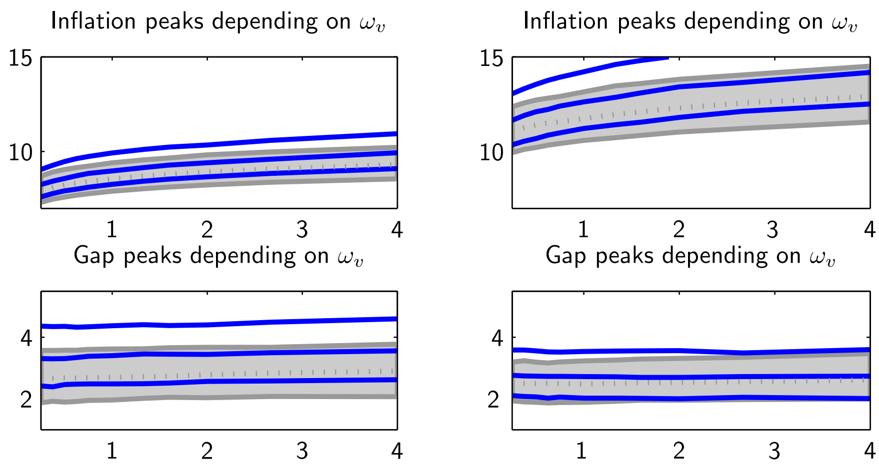

Figure 2.13: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different weights on interest rate smoothing

Results for varying degrees of the interest rate smoothing component. Since with higher values of the interest smoothing weight the policymaker is more reluctant to use its instrument to fight inflation the resulting inflation peaks are increasing in this parameter, as well as the time until inflation is brought back to target, cf. figures 2.13 and 2.15. Besides that the remaining results are not different to those in the benchmark simulation: the highest deviation of unemployment from the Nairu occurs a few quarters after the inflation peak, after one to one and a half year later in the US model and about a year later in the UK model. The perceived values of inflation per-

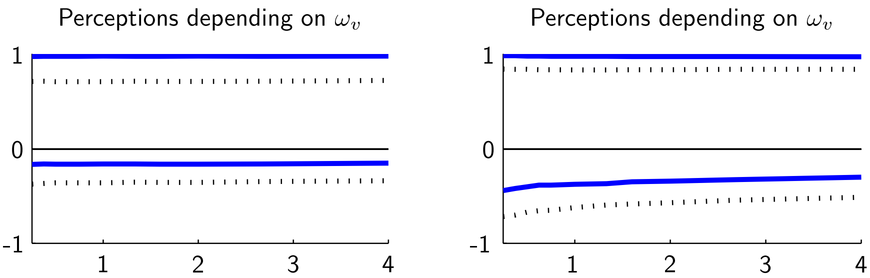

Figure 2.14: Policymakers' perceptions of inflation persistence and Phillips curve slope before and at the height of the GI for different weights on interest rate smoothing

sistence increase from below 0.8 in the periods before the GI to unity at the time of the GI, and the slope estimates are reduced (in absolute terms) to low values $\theta(1) \approx -0.1$ (US) and $\theta(1) \approx -0.5$ (UK) from previously more effective levels, $\theta(1) \approx -0.4$ (US) and $\theta(1) \approx -0.7$ (UK). This pattern is roughly the same for all smoothing weights under consideration. Not surprisingly, although the probability of a GI occurring is high for all values, it is even higher for a relatively greater weight on the smoothing component.

Sensitivity of results for different degrees of inflation bias. Our benchmark used a value close to one thus almost eliminating the desire to attain unusually low unemployment rates. In this section we consider a grid over the interval $[0.5, 1]$ with step size 0.1, conduct $n = 1,000$ simulations for each point, and present the quartiles of the key statistics in the subsequent figures.

The economic intuition of the inflation bias parameter κ is that the policymaker acts in such a way that economic outcomes feature unnecessarily high inflation with no gain in real performance. Consequently, for low values we would expect inflation peaking at higher levels with the unemployment gap peaks remaining unaffected. The first part of this presumption is confirmed by figure 2.16. Indeed, across all variations considered in this section, the inflation bias parameter κ has the greatest effect on inflation outcomes. However,

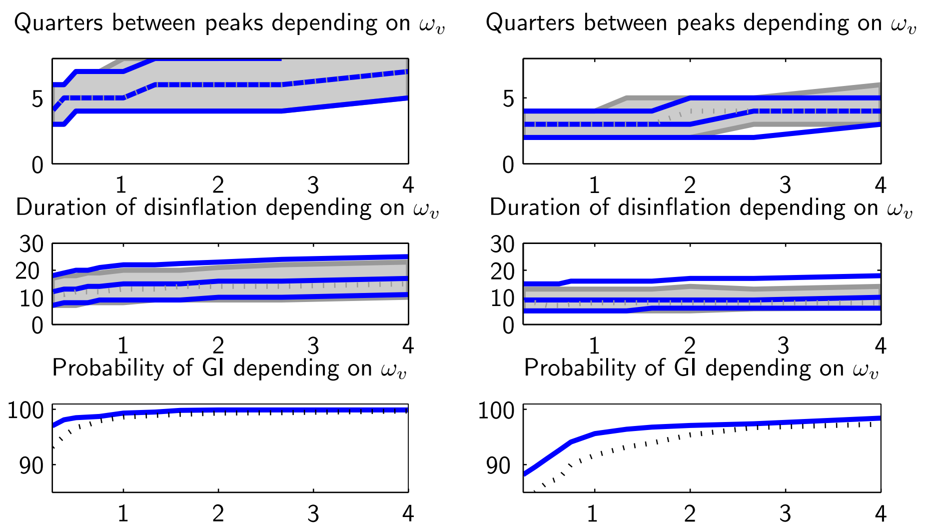

Figure 2.15: Periods that unemployment peak lags inflation peaks (top panel), duration of disinflation and probability of a GI for different weights on interest rate smoothing

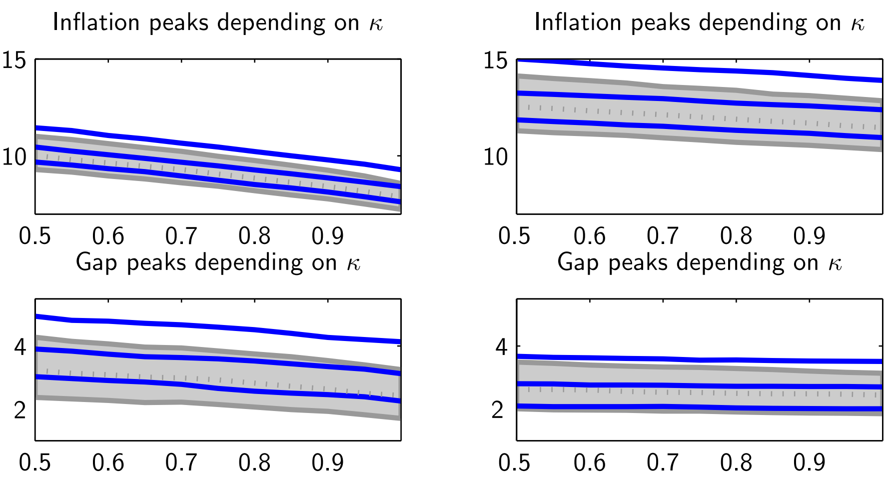

Figure 2.16: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different degrees of inflation bias

while according to our intuition the UK gap is unaffected by changes in κ, US unemployment is brought to higher levels relative to the Nairu for low values of κ. Overall, the effect of κ on the gap is less pronounced than on inflation. A striking difference is the substantially prolonged disinflation period in the US model. This, too, does not come unexpected, as a policymaker focussing

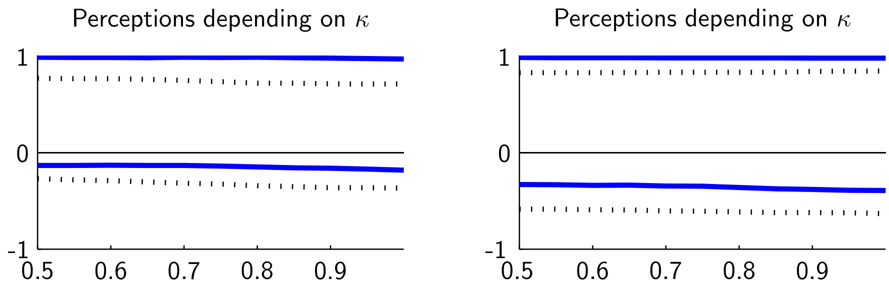

Figure 2.17: Policymakers' perceptions of inflation persistence and Phillips curve slope before and at the height of the GI for different degrees of inflation bias

on the real part of the economy will tend to require more quarters to disinflate more smoothly.

Finally it should be noted that for a high inflation bias (low κ) the probability of a GI reaches 100 percent, that is the more a central bank is tempted to push unemployment below the Nairu the more likely it induces unusually high inflation if it is simultaneously learning about the dynamics of the economy. For κ approaching one, the probability of unusually high inflation rates rapidly decreases, but still remains high for a learning policymaker without the temptation to surprise inflate the economy.

This temptation seems to be absent in our models anyways. The reason we included the inflation bias parameter κ was to account for a Barro and Gordon (1983)–type motivation of the policymaker to deliver surprises in the inflation rate which would result in below average unemployment rates if agents would not erode this temptation by adjusting their expectations accordingly. Hence, we would expect the unemployment gap that the policymaker wishes to reduce to lead inflation which would be the tool used to reach this goal in this model. Our results, however, indicate that inflation is the problem that is solved by pushing unemployment above the Nairu, thus resulting in gap peaks lagging inflation, as confirmed by figure 2.18.

Considering different learning algorithms. The gain value reflects two opposing forces: one is the desire to put higher weights on more recent observations to improve estimates in presence of structural breaks; in this regard higher gain values perform better, given that the learning dynamics are

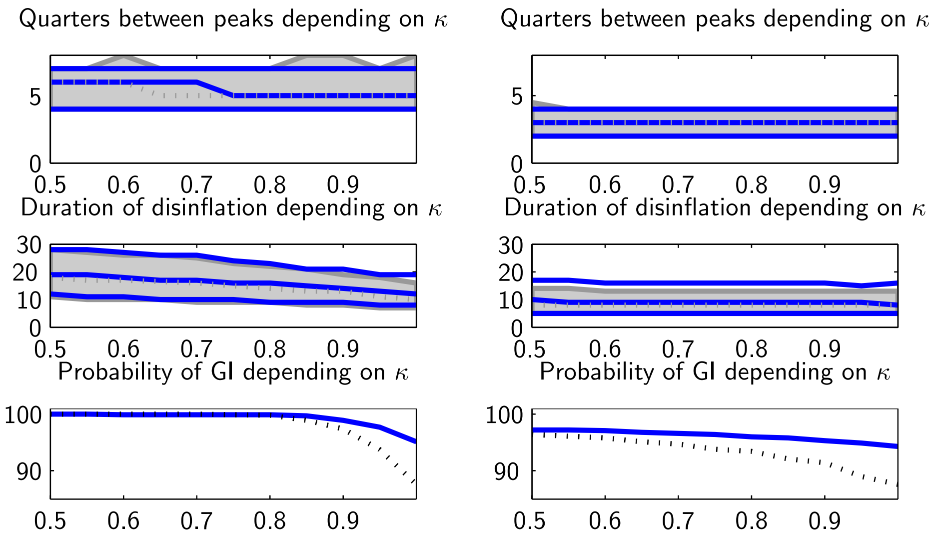

Figure 2.18: Periods that unemployment peak lags inflation peaks (top panel), duration of disinflation and probability of a GI for different degrees of inflation bias

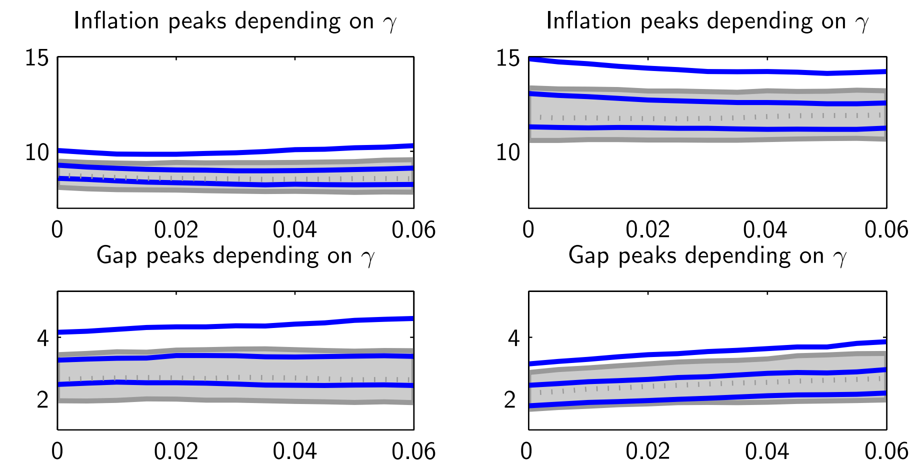

Figure 2.19: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different learning gain parameters γ

substantial as discussed earlier; the other is the desire to use sufficient sample information; if the gain value is too high there are effectively too few observations included in the estimates. The accord in the learning literature is on a gain value of $\gamma = 0.03$, but values around this particular value are equally reasonable. Specifically, we allow for gain values in the range $[0, 0.06]$. It should be noted, that with our non–recursive, discounted least squares formulation, a zero gain coincides with ordinary least squares, cf. appendix A.1, so our analysis nests all relevant learning specifications. While the economic outcomes

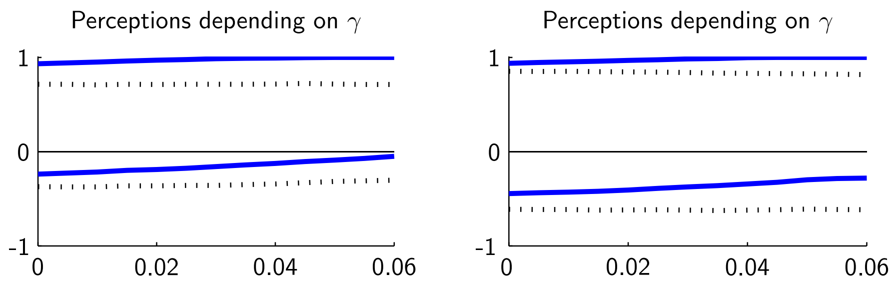

Figure 2.20: Policymakers' perceptions of inflation persistence and Phillips curve slope before and at the height of the GI for different learning gain parameters γ

(inflation and gap peaks) are largely unaffected by the gain value, the perceptions of inflation persistence and Phillips curve slope are more pronounced for higher gain parameters. This reflects the fact that the economic conditions surrounding the height of the GI are causal for the changes in these key parameters and higher gain values place a higher weight on these more recent events. Since this favors the emergence of the GI, its probability slightly increases in the gain parameter.

More importantly though, changes in the learning algorithm has no notable impact on our findings. It is thus not the particular way we model learning, but the accounting for imperfect knowledge as such that drives our results.

Robustness versus different assumptions on Nairu smoothness. We choose a value for filter parameter the policymaker uses to extract the Nairu that reproduces the smoothness in the Nairu as documented in empirical studies. However, our insights should not depend crucially on this choice, so it is important to assess its relevance for our findings. In figures 2.22 – 2.24 we plot our key results against an exponent c, such that the HP parameter is given by $\mu = 10^c$. The smallest value under consideration, $c = 3.2$, is such that it implies the standard calibration $\mu = 1600$, but we include higher values which imply Nairu paths comparable to those typically reported in the related literature.

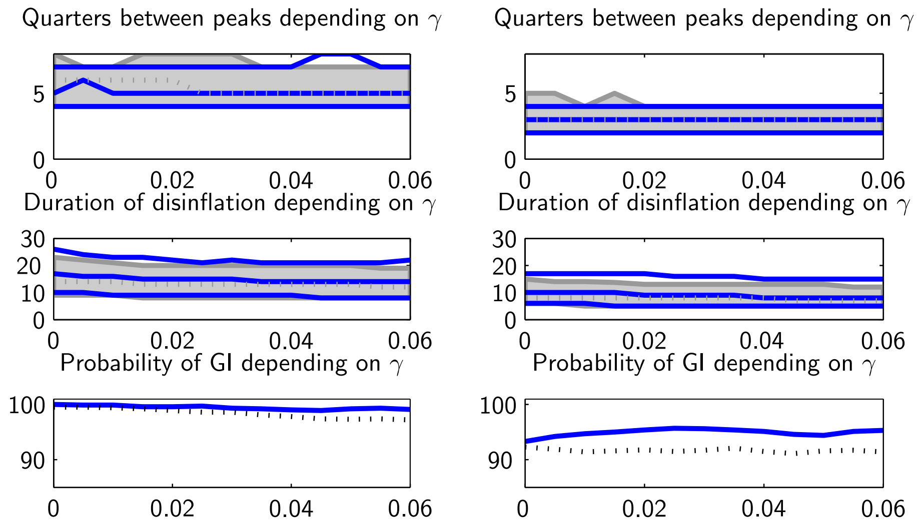

Figure 2.21: Periods that unemployment peak lags inflation peaks (top panel), duration of disinflation and probability of a GI for different learning gains γ

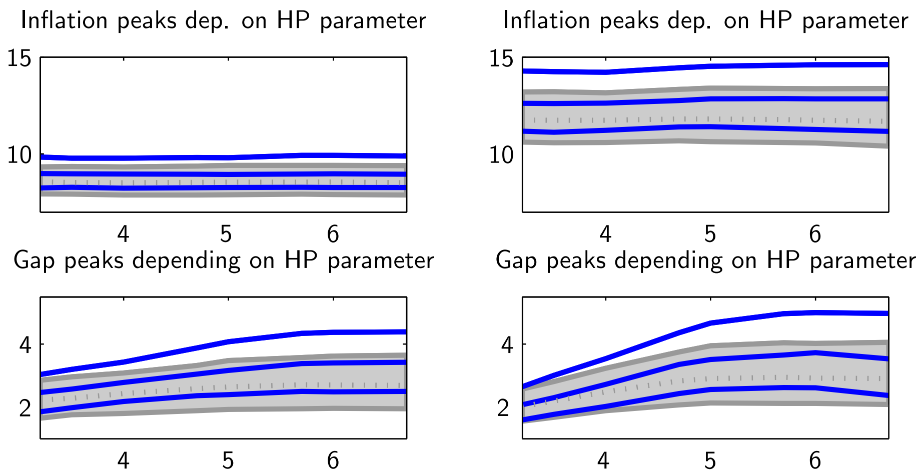

Figure 2.22: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different HP Filter smoothing parameters

The results with different HP parameters μ yield useful insights in the mechanism regarding the Nairu misperceptions. A high smoothing parameter implies a rather flat path for the Nairu. As the gap is determined as the difference of unemployment to the Nairu, the fluctuations in unemployment are soaked up by the perceived gap. Conversely, a low μ allows the Nairu to take much of the variation in unemployment, leaving little movement left for the gap. Thus, higher smoothing parameter imply gap estimates with higher amplitudes. What eventually matters is the relation of amplitudes in the perceived and the true gap. If for a given choice of μ the policymaker would obtain a good approximation to the true gap, then increases in the smoothness parameter would result in excessively high amplitudes in the perceived gap. With the true gap unaffected the policymaker would believe that changes in the gap, though of big size, have not that big of an effect. This effect is stronger

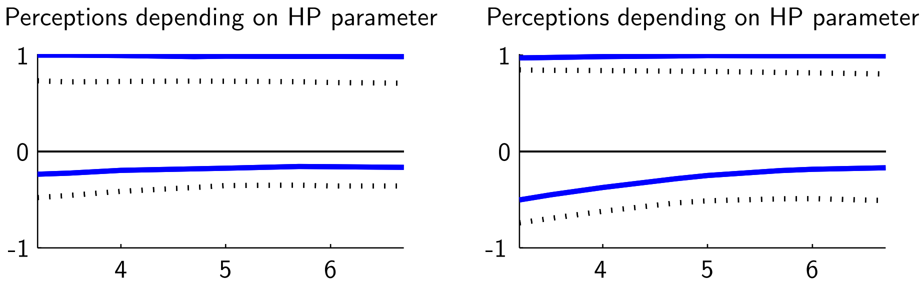

Figure 2.23: Policymakers' perceptions of inflation persistence and Phillips curve slope before and at the height of the GI for different HP Filter smoothing parameters

for the UK model where the true Nairu fluctuations are higher compared to the US model, Gordon (1997) and Greenslade et al. (2003). A low HP smoothing value in the UK model extracts a gap whose fluctuations are small in size and hence agree with the true gap. The reason that this effect appears hardly present in the US simulation is that the true Nairu is relatively flat and thus the gap has a high amplitude. Either the policymaker uses also a high μ yielding gap estimates corresponding to the true gap, or he chooses low values yielding an estimated gap path being smaller in amplitude than the true gap, but then he would observe small changes in the gap exerting already substantial effect on inflation.

This is reflected in the graph plotting the perceived slope against the HP parameter. The US slope estimates are increasing, approaching zero from below, in the HP parameter but not as strongly as in the UK case. Besides that, both models display the same behavior, with the realized maximum gap values admitted by the policymaker to disinflate being higher when the time variation in the gap is considered high, i.e. for high μ. The inflation peaks are little affected by this choice, as are the perceptions of inflation persistence, the speed of the disinflation and the probability of an inflation outbreak. Hence the insights gained in the benchmark simulations carry over to arbitrary values of the filter parameter.

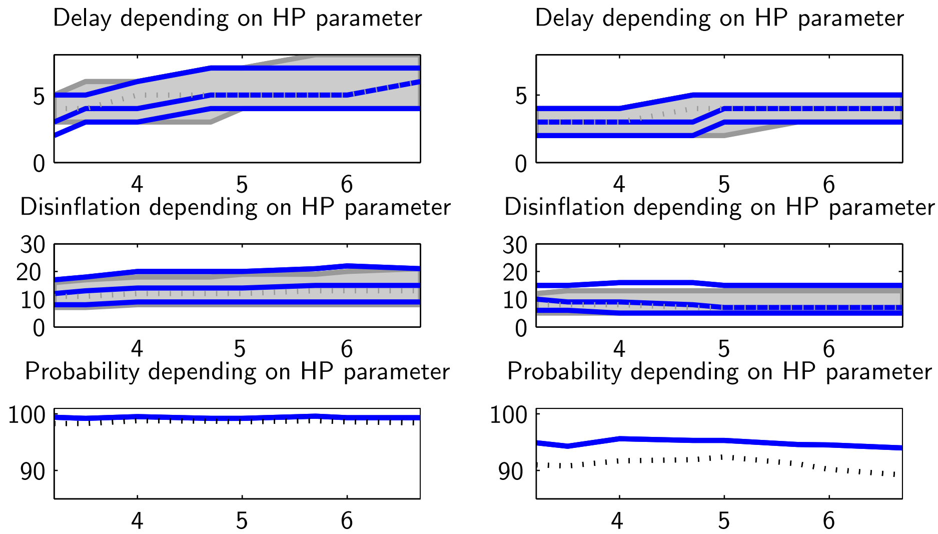

Figure 2.24: Inflation peaks (top panel) and Peaks of the Unemployment Gap (bottom panel) for different HP Filter smoothing parameters

2.6. Learning with a conservative policymaker

We have seen that policymakers with restricted information sets will almost inevitably face inflationary outbreaks as experienced during the GI. However, while optimizing policy leads to unusually high levels of inflation, it does rarely produce outcomes as bad as those experienced in the UK, and in part this holds true for the US case as well.

In our companion paper we argued that the uncertainty of the estimates suggests a Brainard (1967) type moderate policy stance, which implies the muted policy setting we documented in chapter 1. As argued by Brainard and Blinder (1997) high standard deviations of key parameters induce the decision maker to compute the optimal response under certainty and then implement only a fraction of it, a feature Blinder coined policy conservatism.

We thus examine the standard deviations of the policymakers' estimates, depicted in figure 2.25. As can be seen, during the great inflation the uncertainty was substantially higher than in the subsequent periods. Notably the Phillips slope estimates until the early 1980s are very imprecise, a fact that is particularly pronounced in the UK model, thus calling for an especially mute

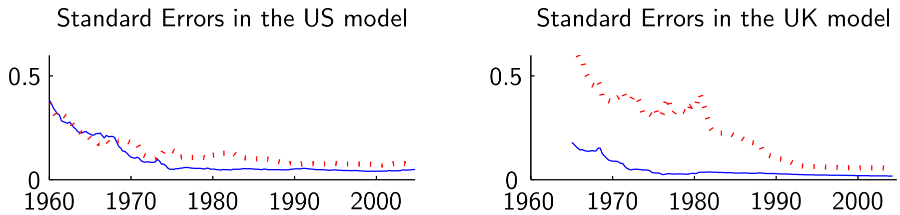

Figure 2.25: Uncertainty of Beliefs in the US (left panel) and UK model (right). Measured by the standard deviations of inflation persistence and of the Phillips curve slope (dashed line).

interest response. This matches the finding in our companion paper, that documents highly positive comovement of historical and optimal interest rates but with a magnitude being persistently below recommended values.

Our modelling strategy was to derive optimal policy decisions of a learning central banker, but we followed the standard approach that focusses on decision making under certainty equivalence. The crucial feature of adaptive learning is the time variation that is due to perpetual refinement of model estimates, but these estimates are surrounded by uncertainty. In the same way the estimates might change, converging to true underlying values, the associated uncertainty will also change, probably decrease over time. Thus, a learning policymaker that takes multiplicative uncertainty à la Brainard (1967) into account will behave conservative, but less and less so as he eventually learns the true parameter values and becomes more and more confident about them. To assess whether a conservative optimizing policymaker would reproduce the bad historical outcomes, we construct a simulation exercise that accounts for the observed phenomenon.

Sack (2000) analyzes Brainard's ideas in a dynamic, infinite horizon model. He proposes a method that allows a reformulation of the decision problem under multiplicative uncertainty in terms of a certainty equivalence problem with a modified loss function. He expresses his results in general terms allowing for a constant in the transition law and the loss functions, but since the constant can always be included in the state vector as in our example, we can simplify his terms determining the optimal instrument setting. In the setup of our optimal linear regulator problem, as discussed in appendices A.2 and A.3, he demonstrates that under certainty as well as under uncertainty, the optimal instrument i is set as

$$i_t = -(B'\Lambda B)^{-1}B'\Lambda A y_t. \tag{2.10}$$

The only difference between both cases is that the matrix Λ solves a different Lyapunov equation. In the certainty case the solution to this equation depends on the weighting matrix in the policymakers loss function, which we labelled Q^*. Sacks result now reduces to the finding that the solution under uncertainty given a weighting scheme Q^* denoted by $\Lambda_u(Q^*)$, relates to the certainty equivalence solution as

$$\Lambda_u(Q^*) = \Lambda(Q^* + \Sigma^*). \tag{2.11}$$

That is, one obtains the solution matrix for the uncertainty case by computing the certainty equivalence solution for a problem where a matrix Σ^* adds to the loss function's weighting matrix. Σ^* collects the covariances between the uncertain parameters. As Blinder argues, Brainard's analysis focusses on the special case with zero covariances between different parameters, so Σ^* would be diagonal. However, Sack's result then suggests that in his framework Brainard type uncertainty can be accounted for by modifying the weights in the loss function. But changes in these weights will alter the way the decision maker balances his target variables and thus the whole paths of them. Except in special cases this will result in policies that are not necessarily correlated and scaled down as in Brainard's static setup. Sack's dynamic analysis thus yields more general results, and for that matter, we already analyzed the effect of uncertainty implicitly, by investigating changes to the loss function's weight in the previous section.

Instead, we proceed by stipulating that policymaker's were intrigued by Brainard's recommendation, which just came up at the time of the great inflation, and decided to account for the high uncertainty of their estimates by muting their response as suggested by his considerations. We model this by computing the optimal interest rate response and then scaling it down, just as a Brainard type policymaker would do. We allow this down scaling for the historical great inflation periods, which we limit to the 1970s, and choose a scale factor given by the ratio of historical to optimal interest rates as documented in chapter 1 (cf. table 1.1 there).

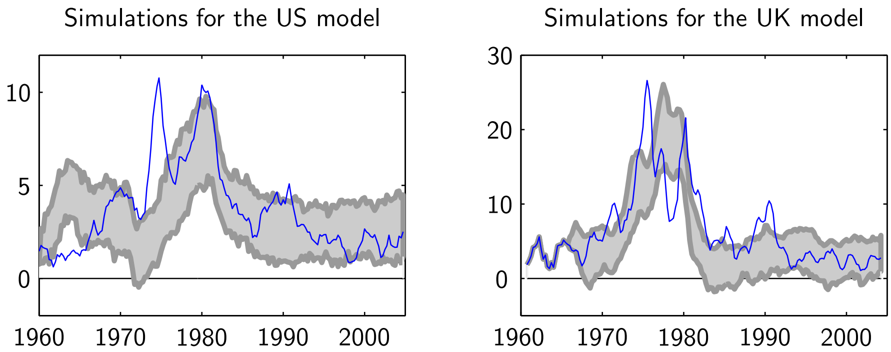

Figure 2.26: Inflation rates induced by a central banker constrained to behave conservative in face of high estimation uncertainty.

We draw $n = 200$ random shock sequences and use our benchmark policy calibration to generate interest rate paths of a conservative policymaker under imperfect knowledge who is otherwise behaving optimally. Figure 2.26 depicts the simulated paths of inflation and contrasts them with the historical time series. Apart from the first peak in the inflation rate, which is commonly attributed to the oil price shock, and therefore to the bad luck factor that we ruled out in our simulations, the similarity of simulated paths to historical inflation rates is striking. The simulations do not only capture the expanded increase in inflation, the extend of its duration, but they yield inflation peaks almost coinciding with observed data. The simulations consistently point at an inflationary outbreak of exactly the dimension, duration and extend experienced in the US and the UK. Furthermore, the disinflation is as rapid as documented in empirical accounts. We thus conclude that conservative behavior plays a crucial role, and should be accounted for in the empirical analysis of models under adaptive learning.

2.7. Conclusions

We pursued the question whether the Great Inflation as experienced in the US and UK can be reconciled with optimal policy under imperfect information. We started from initial beliefs determined from data prior to 1960 and required the central banker to jointly learn about the parameters of a reduced form New Keynesian model and to infer estimates about the unobservable unemployment gap. We saw that this inevitably pushes him into a combination of substantially underestimated levels of the Nairu, very low values of inflation persistence and favorable slope coefficients of the Phillips curve.

The initial rise in inflation thus calls for no immediate reaction. Inflation is not persistent, hence any shock to it will fade out sufficiently fast. Unemployment, however, is perceived to be best held at low levels, which in retrospect unfolded inflationary pressures without observers identifying this source, and low deviations from the Nairu appear effective enough to counteract inflation.

As inflation reaches high levels, agents realize its persistent nature and correspondingly conclude that tight inflation control is in order. However, since the misperception in the gap variable–the real time perception of the gap was substantially positive while the true gap was close to zero–falsely lead agents to the insight that even substantial gap values do not release tension in the inflation rate. Thus at the outbreak of the great inflation, interest movements appear to have a limited effect.

As soon as sufficient movement in unemployment allows for a more precise detection of the Nairu and a clearer assessment of the Phillips curve slope, the policymaker pushes unemployment over the true Nairu until inflation quickly returns to moderate values. At this point the unemployment gap is also allowed to fade out. The simulations confirm the duration of the disinflation period and the lags that are necessary for unemployment to exceed the Nairu.

These evolutions in beliefs and the associated stylized facts as summarized above are consistently reproduced in simulations of the system under learning. Our finding is true for any type of optimizing policymaker from a wide range of preferences and it applies equally well to the US and the UK. The difference in both models is the substantially higher uncertainty surrounding the slope coefficient in the UK model, inducing the policymaker to act more conservative. Indeed, modelling a Brainard type central banker that implements only a fraction, that is decreasing in uncertainty, of optimal interest rates produces outcomes almost identical to the great inflation episodes in both countries, being comparable in the duration of the periods of increasing and subsequent sustained high inflation, the rapid disinflation, as well as the dimension of the high rates experienced in both countries.

We therefore find support for the notion of adaptive learning in inflation unemployment dynamics, and provide a better account for one of the most important periods in US and UK postwar inflation history. Our considerations also point to a substantial role to uncertainty in empirical models of adaptive learning, which we leave as a future research project.

CHAPTER 3

Estimating DSGE models under Adaptive Learning

3.1. Introduction

Over the last decade, microfounded dynamic stochastic equilibrium (DSGE) models have become the main tool of macroeconomic analysis. The New Keynesian model analyzed in Woodford's (2003) monograph and in Clarida et al.'s (2000) prominent contribution, too name only a few, is one of the most important contributions to monetary economics. Consequently, many authors estimate versions of these models, usually based on the rational expectations (RE) hypothesis. While this approach is very fruitful there has recently been an important branch of research that replaces this informationally demanding approach with adaptive learning (AL), where agents still act optimally but face imperfect knowledge of the economic structure. Ireland (2003) calls for irrational expectations econometrics that account for restricted information sets. As a response to this call, this paper offers a general approach to analyze DSGE models under adaptive learning, enabling researchers to contrast corresponding results with the rational expectations benchmark. This method extends the standard estimation procedure by incorporating results from Evans and Honkapohja (2001) into the analysis. The second part of the paper applies this procedure to Ireland's (2004) assessment of the New Keynesian model amended with real business type technology shocks. Estimation of his model under adaptive learning gives a better fit to US data then the RE benchmark and yields substantially different insights.

3.2. The basic approach

Microfounded models incorporating dynamic and forward looking elements have become the standard tool of macroeconomic analysis. The associated first order as well as the market clearing conditions and relations describing preferences and technologies give rise to a set of equations that must be satisfied in equilibrium. These equations link current and past realizations of endogenous variables to their expected future values and exogenous shocks. Formally, we can write such a dynamic model of the economy as

$$E_t F(\tilde{y}_{t+1}, \tilde{y}_t, \tilde{y}_{t-1}, \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}; \Psi) = 0 \tag{3.1}$$

where E_t is the expectations operator conditional on date–t information, the mapping F summarizes the model equations involving the endogenous vector $\tilde{y}$ and the exogenous terms $\tilde{\varepsilon}$. The vector Ψ collects all parameters of the model. To analyze such an economy the mapping is typically log–linearized around the deterministic steady state, which is implicitly defined by the rest point of the system $\bar{y}$ given no more shocks are ever to occur,

$$F(\bar{y}, \bar{y}, \bar{y}, 0, 0; \Psi) = 0 \tag{3.2}$$

and is reformulated in terms of log deviations from steady state values, $y_t = \log(\tilde{y}_t) - \log(\bar{y})$ and $\varepsilon_t = \log(\tilde{\varepsilon}_t)$

Log linearizing the mapping F yields

$$F \simeq \text{Jac}(F)\text{diag}(\bar{y}, \bar{y}, \bar{y}, \bar{\xi})(y_{t+1}, y_t, y_{t-1}, \varepsilon_t, \varepsilon_{t-1})' \tag{3.3}$$

Together with (3.1) this yields the canonical linear structural form

$$y_t = \Gamma_1 E_t y_{t+1} + \Gamma_2 y_{t-1} + \Gamma_3 \varepsilon_t \tag{3.4}$$

with the shock vector having possibly the autoregressive representation

$$\varepsilon_t = \Lambda \varepsilon_{t-1} + \xi_t \tag{3.5}$$

In this formulation, all matrix entries are functions of the underlying deep parameters Ψ, as captured by the Jacobian matrix Jac(F) and the steady state values given by the diagonal matrix appearing in (3.3). The vector ξ is serially and intertemporally uncorrelated white noise, i.e. $E[\xi_{j,t}\xi_{k,s}] = \sigma_j^2$ only for $j = k$ and $t = s$ and zero otherwise.

Standard solution methods, e.g. Sims (2002), Klein (2000) or Uhlig (1995), yield a reduced form representation

$$y_t = \Theta_1 y_{t-1} + \Theta_2 \varepsilon_t. \tag{3.6}$$

This representation along with the noise equation is subsequently mapped into an appropriate state space form, for which the likelihood can be evaluated using the standard Kalman Filter. The underlying structural parameters Ψ can then be estimated by the maximum likelihood method.

The mapping to a state space form is commonly necessary due to the presence of unobservable variables. Hence we will consider this mapping onto a general state space form

$$m_t = Ms_t + X_m x_t + M_\zeta \zeta \tag{3.7a}$$

$$s_t = Ss_{t-1} + X_s x_t + S_\zeta \zeta \tag{3.7b}$$

where it is important to note that all matrices are functions of Ψ. The Kalman Filter computes a set of recursive equations in the one period ahead forecast of the measurement and state vectors m and s and in the corresponding forecast errors' covariance matrices, Υ and Σ, which can be used to evaluate the likelihood associated with the underlying parameter vector.

To summarize, for a given parameter vector Ψ we compute the linearized structural form $(\Gamma_1, \Gamma_2, \Gamma_3, \Lambda)$ which by standard RE reasoning implies an equilibrium representation $(\Theta_1, \Theta_2, \Lambda)$ as given by equation 3.6. To estimate this latter system, it is mapped onto a time invariant state space form (3.7), mainly due to the need to distinguish observable from unobservable variables. Each vector Ψ thus induces a unique time invariant state space form and with it a unique likelihood value. Standard optimization routines[1] can then be used to infer the maximum likelihood estimate Ψ_T^* of the structural parameters from a sample of length T.

3.3. Agents' beliefs and implied dynamics

To extend the basic procedure to deal also with the adaptive learning hypothesis, we introduce the concept of the T–map relating the perceived law of motion (PLM) and the actual law of motion (ALM), as defined by Evans and Honkapohja (2001).

The PLM captures the model dynamics as perceived by the agents. It is clear, that many variations are possible here, but the standard approach as suggested by Evans and Honkapohja (2001) is to endow the agents with a model of the same functional form as the equilibrium representation (3.6),

[1] Sims's (2002) Matlab code *csminwel* offers an appropriate and commonly applied algorithm.

thus the PLM takes the recursive form

$$y_t = \Omega_1 y_{t-1} + \Omega_2 \varepsilon_t \tag{3.8}$$

$$\varepsilon_t = \Omega_3 \varepsilon_{t-1} + \xi_t \tag{3.9}$$

with Ω_1, Ω_2 and Ω_3 not necessarily coinciding with their rational expectations equilibrium counterparts Θ_1, Θ_2 and Λ. Given a PLM of this form, agents forecasting model is given by

$$E_t y_{t+1} = \Omega_1^2 y_{t-1} + (\Omega_1\Omega_2 + \Omega_2\Omega_3)\varepsilon_t \tag{3.10}$$

The underlying structural form (3.4) maps the agents' forecast model into an actual law of motion

$$y_t = (\Gamma_1\Omega_1^2 + \Gamma_2)y_{t-1} + (\Gamma_1(\Omega_1\Omega_2 + \Omega_2\Omega_3) + \Gamma_3)\varepsilon_t \tag{3.11}$$

That is the beliefs of the agents, parameterized by the matrices $(\Omega_1, \Omega_2, \Omega_3)$ are mapped onto actual transition matrices. The mapping is usually referred to as the T–map, and is given in our terminology by

$$T : \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix} \mapsto \begin{pmatrix} \Gamma_1\Omega_1^2 + \Gamma_2 \\ \Gamma_1(\Omega_1\Omega_2 + \Omega_2\Omega_3) + \Gamma_3 \\ \Lambda \end{pmatrix} \tag{3.12}$$

The consistency condition under rational expectations where agents have perfect information requires their forecast model (PLM) to coincide with the induced dynamics of the model (ALM), hence the RE equilibrium can be characterized as the fixed point of the T–map. Solution algorithms as suggested by Blanchard and Kahn (1980) and extended by Klein (2000), Sims (2002) and others solve this fixed point problem.

For completeness we note that many studies in the AL literature focus on whether agents who initially use a PLM different from the REE fixed point will ever learn the equilibrium parameters despite the fact that their misperceived forecast model crucially feeds back into the dynamics of the model. Given that they use consistent estimators such as OLS, Evans and Honkapohja (2001) show that a necessary and sufficient condition is the stability of an associated continuous time differential equation depending solely on the T–mapping. This boils down to inspecting the eigenvalues of $\text{Jac}(T)$ and verifying whether these have real parts less than unity. This is a standard condition in the theory of ordinary differential equations, and yields equations which can be checked numerically or in simple cases even analytically. Thus, the T–map serves in determining whether the additional dynamics introduced by learning will eventually fade out bringing the economy into its full information rational expectations equilibrium. However, we will use the T–map to track the feedback introduced by learning agents, so we will be able to use data series and estimate a DSGE model under the assumption of adaptively learning agents.

3.4. Estimation of PLM

We shall first discuss how estimation proceeds given a sequence of PLMs that agents have estimated over time, while discussing later how this sequence is actually generated. So suppose that since date t_0 agents have at each date reestimated their model of the economy in order to enable forecast of relevant variables. This yields a sequence of perceived models, $\{\Omega_{1,t}, \Omega_{2,t}, \Omega_{3,t}\}_{t=t_0}^{T}$. Since agents base their forecasts $\hat{E}_t y_{t+1}$ for each date t on the corresponding estimates, the actual dynamic evolution of the economy at date t is parameterized by $T(\Omega_{1,t}, \Omega_{2,t}, \Omega_{3,t})$.

VAR approach. The vector autoregression (VAR) approach simply assumes that agents use a VAR model as PLM (3.8). This means that standard VAR estimation yields estimates for the PLM matrices. The VAR approach has the advantage that dynamics present in the data but not captured by our stylized model are taken account of. Furthermore, it is a simple approach that does not require deep understanding of the underlying model. Unfortunately, the VAR method requires all variables to be observable, which in general is not true. So agents might want to define appropriate proxies for unobservable variables prior to estimation, e.g. detrending real output to obtain a measure of the output gap.

Model consistent approach. As an alternative we might assume that agents impose more structure on their forecast model than in a simple VAR. While the VAR approach only requires agents to behave individually optimal and tackling their forecast problem in a convenient but simplifying way, one might also stipulate that they know the structure of the nonlinear aggregate dynamic relations, though not its exact parametric specification. Thus we might as well assume that in each period they perform the full DSGE estimation under rational expectations, which accounts for all cross equation restrictions imposed by the model. This yields a sequence of deep parameters that evolve over time, and with them a sequence of forecast models, derived via the relations described above, i.e. agents plug the deep parameter estimates into the model equations, log–linearize and solve the structural model for a reduced form difference equation that they can use to form forecasts. Since this approach respects the model equations and embodies learning only insofar as it reflects the time variation in parameter estimates, we label this procedure the model consistent approach.

3.5. Application to Ireland's model

Ireland (2004) develops a New Keynesian model which explicitly accounts for real business type technology shocks. The goal is to assess its importance relative to the three types of shocks which are usually investigated in this framework, a preference shock, a cost push shock and a monetary policy shock. The model is mapped onto an econometrically usable form and estimated for US data. For convenience, we briefly review the key building blocks of the model and discuss its main difference to the standard model.

The model. Ireland's (2004) model features a representative household who maximizes a discounted, infinite stream of single–period money–in–the utility preferences

$$E\sum_{t\geq 0}\beta^t[a_t \log C_t + \log M_t/P_t - h_t^\eta/\eta] \tag{3.13}$$

subject to the budget constraint

$$M_{t-1} + B_{t-1} + T_t + W_t h_t + D_t \geq P_t C_t + B_t/r_t + M_t. \tag{3.14}$$

Finished good consumption C_t and real money balances M_t/P_t enter as logarithms while hours work h_t have an elasticity of $\omega \equiv 1/\eta$. Agents allocate their nominal wealth consisting of their wage receipts $W_t h_t$ and their stock of previously accumulated money and bonds $M_{t-1} + B_{t-1}$ to consumption $P_t C_t$ and a portfolio of money M_t and bonds worth B_t in the next period given an interest rate r_t. One of the first–order conditions of the agent's problem is given by

$$h_t^{\eta-1} = \frac{a_t}{C_t}\frac{W_t}{P_t}. \tag{3.15}$$

The variable a_t is a log–AR(1) preference shock, that makes consumption more attractive. A positive realization of this shock induces agents to work more in order to increase wage income for consumption purposes. The responsiveness of labor supply increases with the elasticity parameter ω. A further first–order condition relates current and future consumption,

$$\frac{a_t}{C_t} = \beta E_t\Big(\frac{a_{t+1}}{C_{t+1}}\frac{P_t}{P_{t+1}}\Big) \tag{3.16}$$

that is if e.g. the expected future preference shock exceeds the current shock, consumption is shifted from today towards tomorrow. The preference shock thus impacts on the intertemporal relation of the consumption stream (NO, cancels out if AR1!) as well as the intratemporal consumption–labor decision, where the latter is governed by the elasticity of labor ω. It will turn out that the latter parameter and the size of this shock play a major role in US data.

There is a competitive final goods sector facing a constant returns to scale technology

$$Y_t \leq \Big(\int_0^1 Y_t(i)^{(\theta_t-1)/\theta_t}\Big)^{\theta_t/(\theta_t-1)} \tag{3.17}$$

where the parameter θ_t plays the role of a stochastic markup or cost–push shock. The underlying stochastic process is also assumed to be a first–order AR(1). Intermediate goods producers, distributed over the unit interval and indexed by i, are monopolistically competitive firms, also facing a constant returns to scale technology

$$Y_t(i) \leq z_t h_t(i), \quad i \in [0,1] \tag{3.18}$$

where z is an RBC type unit root technology shock. Intermediate goods firms face Rotemberg (1982) price adjustment cost

$$\frac{\phi}{2}\Big(\frac{P_t(i)}{\pi P_{t-1}(i)} - 1\Big)^2 Y_t \tag{3.19}$$

This term punishes changes in individual prices unless the change corresponds to the steady state inflation rate. Therefore the reaction to shocks will be smoothed out over several periods inducing less flexible prices.

The model is closed by specifying that the central bank sets the nominal interest rate r_t by following a version of Taylor's (1993) rule, including output growth g_t (a hat indicates deviations from steady state values)

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + \rho_\pi \hat{\pi}_t + \rho_g \hat{g}_t + \rho_x \hat{x}_t + \varepsilon_{rt} \tag{3.20}$$

Ireland proposes the use of a difference rule, thus fixing the inertial parameter in the rule $\rho_r = 1$. ε_r is a shock to this interest rate rule reflecting non–systemic deviations.

This model differs to the standard New Keynesian model as developed by Woodford (2003) and discussed e.g. by Clarida et al. (2000) in the introduction of the technology shock z_t. The log–linearized first order conditions are given by

$$x_t = \alpha_x x_{t-1} + (1-\alpha_x) E_t x_{t+1} - (r_t - E_t \pi_{t+1}) + (1-\omega)(1-\rho_a) a_t \tag{3.21}$$

$$\pi_t = \beta \alpha_\pi \pi_{t-1} + \beta(1-\alpha_\pi) E_t \pi_{t+1} + \psi x_t - e_t \tag{3.22}$$

$$g_t = y_t - y_{t-1} + z_t \tag{3.23}$$

$$x_t = y_t - \omega a_t \tag{3.24}$$

$$r_t = \rho_r r_{t-1} + \rho_\pi \pi_t + \rho_g g_t + \rho_x x_t + \varepsilon_{rt} \tag{3.25}$$

which using Ireland's notation $s_t = (y_{t-1}, r_{t-1}, \pi_{t-1}, g_{t-1}, x_{t-1}, \pi_t, x_t)'$ can be rewritten in compact matrix notation as

$$C_0 s_t = C_1 s_{t-1} + C_2 E_t s_{t+1} + C_3 v_t \tag{3.26}$$

where the C–matrices collect the parameters of equations (3.21)–(3.25).

$$s_t = \Gamma_1 E_t s_{t+1} + \Gamma_2 s_{t-1} + \Gamma_3 v_t \tag{3.27}$$

with

(3.28)

$$\Gamma_1 = \frac{1}{\Delta}\begin{pmatrix} \rho_g & -\rho_r & -\rho_\pi\beta\alpha_\pi & 0 & \alpha_x \\ -\rho_g & \rho_r & \rho_\pi\beta\alpha_\pi & 0 & \alpha_x(\rho_g+\rho_\pi\psi+\rho_x) \\ \psi\rho_g & -\psi\rho_g & \beta\alpha_\pi(1+\rho_g+\rho_x) & 0 & \psi\alpha_x \\ -1+\rho_\pi\psi+\rho_x & -\rho_r & -\rho_\pi\beta\alpha_\pi & 0 & \alpha_x \\ \rho_g & -\rho_r & -\rho_\pi\beta\alpha_\pi & 0 & \alpha_x \end{pmatrix}$$

(3.29)

$$\Gamma_2 = \frac{1}{\Delta}\begin{pmatrix} 0 & 0 & 1-\rho_\pi\beta(1-\alpha_\pi) & 0 & 1-\alpha_x \\ 0 & 0 & \rho_x+\rho_\pi\beta(1-\alpha_\pi)+\rho_\pi\psi+\rho_g & 0 & (1-\alpha_x)(\rho_\pi\psi+\rho_g+\rho_x) \\ 0 & 0 & \psi+\beta(1+\rho_g+\rho_x)(1-\alpha_\pi) & 0 & \psi(1-\alpha_x) \\ 0 & 0 & 1-\rho_\pi\beta(1-\alpha_\pi) & 0 & 1-\alpha_x \\ 0 & 0 & 1-\rho_\pi\beta(1-\alpha_\pi) & 0 & 1-\alpha_x \end{pmatrix}$$

It is worth noting that economic reasoning attaches positive values to all parameters appearing in the solvability condition, implying that it is always satisfied for economically meaningful systems.

Perceived law of motion. Since some of the key variables in our model are unobservable, we prefer the model consistent forecasts over the simple VAR approach, which would require either observability of all variables or an appropriate approximation of them in a first step. The model consistent approach requires the sequential estimation of the DSGE model over all relevant subsamples. The high dimensionality of the parameter vector requires a sufficiently long sample to infer precise estimates, so as shortest subsample we choose the period 1948:2 − −1980:1. In the terminology of the previous section, we thus choose t_0 = 1980:1 and T = 2003:1. Figure 3.1 illustrates the evolution of agents' beliefs about the deep parameters over time; it contains the estimates plotted versus the end point of the subsample, so e.g. the point $(1985, 3.8 \times 10^{-3})$ in the bottom right panel indicates that the estimate for the monetary policy shock's standard deviation was 0.0038 for the 1948–1985 subsample. As expected, the recursive estimates display substantial time variation so therefore the forecast models that feed back into the model via the expectations terms introduce a high degree of time variation in the reduced form–or actual law of motion's–parameters. One might suspect that this accounts at least for part of the instability results documented by Ireland over the two subsamples presented in his paper. The perceived law of motion of any

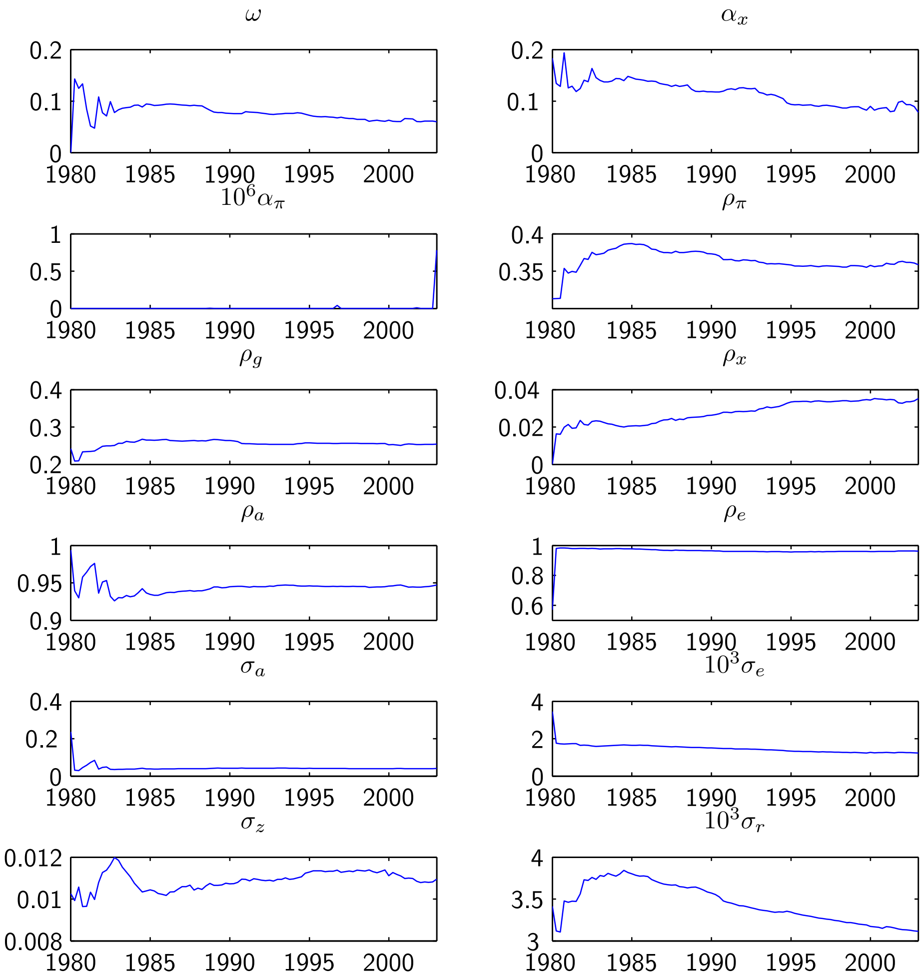

Figure 3.1: Perceived deep parameters over time

given date t is then obtained by plugging the corresponding estimates into the log–linearized system equations (3.27) and solving this structural form equation for the implied rational expectations equilibrium using Klein's method. This yields a dynamic system of the form (3.8), which agents would use as real–time forecast model for the given date. This sequence of time–varying PLMs is entirely determined by the data and will not be influenced by the following estimation procedure.

Estimating Ireland's model under Adaptive Learning. For any parameter vector Ψ, equations (3.27) give a structural form that maps agents beliefs onto actual dynamic outcomes as indicated by the T–map in (3.10). We do not want to determine the fixed point, however, since we do not assume agents to use the forecast model consistent with the parameter vector under consideration, but instead use their real–time perceptions given by the sequence of PLMs. The T–map translates these perceptions onto actual dynamic coefficients. It should be noted that the T–map depends on the structure parameterized by the chosen vector Ψ, and hence for each such vector maps the PLM sequence onto a time–varying actual law of motion. The likelihood of this latter system can be evaluated using the time–dependent version of the Kalman filter, cf. Harvey (2003). We then estimate the deep parameters under adaptive learning by computing the maximizer Ψ^* of this likelihood function.

Estimation Results. The two left columns of table 3.1 contrasts the estimation results for the model under rational expectations with the model under adaptive learning. We will allow the interest rate smoothing parameter ρ_r in the Taylor rule to be estimated from the data while Ireland proposes the use of a difference rule, thus fixing $\rho_r = 1$.

We find that in the RE version of our model, the estimate of the inertial parameter is indeed close to unity. Reestimating the model with this restriction imposed yields a maximized log–likelihood value being only 0.2 below the unrestricted value, thus a likelihood ratio test would not reject the null hypothesis of this parameter being unity. The AL estimate, however, is substantially below one, and the difference in log–likelihood values is 6.1 which implies rejection of the null hypothesis at any standard confidence level. Hence, our first insight from the AL estimation is that we should not fix the inertial Taylor rule parameter, but estimate it instead, which gives us a value of $\rho_r = 0.8$.

The estimate of the labor elasticity in our RE estimation is slightly higher than Ireland's, although it remains statistically insignificant. Our AL estimate is zero right away. This suggests, as discussed by Ireland, that labor is very inelastic. The model thus transmits preference shocks exclusively via an increase in demand, without any substantial increase in labor and thus supply.

	Sample 1980:1–2003:1		Sample 1980:1–1991:3	
Parameter	RE Estimates	AL Estimates	RE Estimates	AL Estimates
ω	0.2567	0.0001	0.1724	0.0013
α_x	0.0007	0.9813	0.0044	0.8856***
α_π	0.0000	0.0000	0.0000	0.0000
ρ_π	0.5925***	0.4290***	0.6253***	0.5217***
ρ_g	0.4005***	0.2178***	0.4445***	0.2949***
ρ_x	0.0951**	0.0086***	0.0704**	0.0000
ρ_r	0.9902***	0.7963***	0.9884***	0.8613***
ρ_a	0.9328***	0.9978***	0.9126***	0.9931***
ρ_e	1.0000***	1.0000***	0.9999***	1.0000***
σ_a	0.0224***	0.3257***	0.0386***	0.0000
σ_e	0.0003***	0.0036***	0.0006***	0.0050***
σ_z	0.0049***	0.0082***	0.0002	0.0109***
σ_r	0.0029***	0.0016***	0.0039***	0.0024***
$\mathcal{L}$	1197.8	1214.8	561.7	570.9

Table 3.1: Estimation results. Estimates that are significant at the 1 percent level are marked (***), significance at the 5 and 10 percent levels are marked (**) and (*).

The estimates of the backward looking elements, α_x and α_π, are not significant. Interestingly though, the point estimate of the persistence in the output gap equation is close to unity in the AL case. Thus, despite the high uncertainty surrounding the estimate, it indicates that backward elements are quite important for the dynamics of the output gap once we allow for imperfect information. This contrasts to the RE result which delivers an estimate of virtually zero.

The estimates of the autoregressive coefficients in the equations for the shock processes are very similar across the two models, both suggesting high persistence of the preference and markup shock. The standard deviations of the shock processes noise terms, however, differ, being larger for the preference, markup and technology shock in the AL case at the expense of a lower monetary policy shock variance.

The most important observation is that the model under adaptive learning attains a higher maximized likelihood value than the same model assuming rational expectations. The log–likelihood of the AL model, $\mathcal{L}_{AL,full} = 1214.8$ with the subindex indicating the use of the full sample, exceeds its RE counterpart, $\mathcal{L}_{RE,full} = 1197.8$, by 17 points, which is a substantial increase. Unfortunately, as neither method nests the other as a special case, or could in any way be parameterized, no formal test can be applied to judge the statistical significance of this difference in likelihood.

The first thing to notice is that, again, the likelihood attained for the AL model, $\mathcal{L}_{AL,1} = 570.9$ with the subindex 1 indicating the use of the first half of the sample, exceeds the value of the RE estimation, $\mathcal{L}_{RE,1} = 561.7$. Thus adaptive learning consistently yields a superior fit to the data. The backward looking component in the output gap equation is again very high and unlike in our previous estimation, where it was found to be insignificant with a nevertheless high t–statistic of 1.5, this term is now significant, reassuring our previous finding of the importance of this term.

To formally assess the issue of parameter stability, we also compute the maximized log–likelihood values of both models for the 1991:4–2003:1 subsample. These are given by $\mathcal{L}_{AL,2} = 663.8$ and $\mathcal{L}_{RE,2} = 659.5$. We use Andrews and Fair's (1988) likelihood ratio test which contrasts the log–likelihood over the two unrestricted subsamples with that of the restricted estimation,

$$LR = 2(\mathcal{L}_{\cdot,1} + \mathcal{L}_{\cdot,2} - \mathcal{L}_{\cdot,full}) \tag{3.30}$$

With our results this test statistic values 39.8 and 46.8 for the AL and RE model, respectively. The asymptotic distribution of this statistic is chi–square with degrees of freedom equal to the number of restrictions, that is the number of parameters allowed to vary between subsamples, i.e. our 13 variables contained in Ψ. Although the AL model yields a lower test statistic, favoring the idea of less time variation in the AL model, both tests decisively reject the null hypothesis of parameter stability. Thus even with learning dynamics accounted for by our method, the data still contains evidence for structural change. At this point we should note that under adaptive learning there is another important aspect that should be taken account of: our modelling strategy was to describe monetary policy with an ad hoc rule. In particular, this rule will not change as the policymaker learns about the economy, while agents are allowed to change their forecasting rules accordingly. It appears plausible to model learning as two–sided in the sense that we also allow for policymakers adjusting their behavior in face of changing beliefs on the dynamics of the economy. This would imply that the policy rule used by the central bank would be an optimal response to the real–time estimate of the economic system, and therefore change over time as well. Hence our result points at the need to model learning as a two–sided process being both relevant for keeping track of updated forecasting models by forward looking agents but also for the optimal behavior of policymakers who revise their optimal plans in face of changes in the economy.

Ireland proposes a decomposition of the variances of the endogenous variables into parts attributable to each of our four shocks. The endogenous variables, output growth g, inflation π, the interest rate r and the output gap x, relate to the state vector s via the relation $y_t = Cs_t$ with an appropriate matrix C, while the state equation $s_{t+1} = As_t + B\varepsilon_{t+1}$ implies a solution for the state vector k periods ahead of

$$s_{t+k} = \sum_{j\geq 0} A^j B\varepsilon_{t+k-j}. \tag{3.31}$$

Applying the date–t conditional expectations operator and subtracting the result from the above term yields the k–period forecast error

$$s_{t+k} - E_t s_{t+k} = \sum_{j=0}^{k-1} A^j B\varepsilon_{t+k-j} \tag{3.32}$$

and hence the k–step ahead forecast error variances

$$\Sigma_k^s = E(s_{t+k} - E_t s_{t+k})(s_{t+k} - E_t s_{t+k})' = \sum_{j=0}^{k-1} A^j BVB'A^{j\prime} \tag{3.33}$$

From the state equation it follows that the unconditional variances Σ^s solve the Lyapunov equation

$$\Sigma^s = A\Sigma^s A' + BVB'. \tag{3.34}$$

The variances of the endogenous variables y then follow as $\Sigma_\cdot^d = C\Sigma_\cdot^s C'$. V is the covariance matrix of our four independent shock terms, $V = \text{diag}(\sigma_a^2, \sigma_e^2, \sigma_z^2, \sigma_r^2)$. The variances due to a single shock are computed by replacing this matrix by a diagonal matrix containing only the variance of the corresponding shock. Since all terms are linear, this yields a decomposition into four terms each containing the part of the overall variance due to a particular shock.

Table 3.2 presents the contributions of the corresponding variance terms in the rational expectations model at horizons $k = 1, 4, 8, 12, 20, 40, \infty$ in percentage terms. Output growth is mainly affected by shocks to preferences. About 50 percent are due to this source while the policy shock contributes to about a third of the overall variance and the technology to roughly 16 percent. The effect of all shocks remains approximately constant across all forecast horizons. Inflation is affected at short horizons mainly by the cost–push shock, which initially contributes 58 percent but becomes more important in the long run resulting in being the dominant factor in the unconditional variance. The presence of a near unit root, however, makes the unconditional results difficult to compare, so we also focus on our finite long–term horizon forecast $k = 40$. At short horizons the policy shock and to a lesser extend the technology shocks are also important, injecting 29 percent and 13 percent of the overall one–step ahead variance. These contributions decrease as the long term variances are more and more affected by the cost–push shock, resulting in minor contributions of these two shocks, both effects being exceeded by that of the preference shock which contributes to 7 percent for $k = 40$. The interest rate is affected mainly by the preference shock in the short run, accounting for 85 percent of the overall variance. With an increasing forecast horizon the cost–push shock becomes more and more important, accounting for 47 percent of our long term variance measure. The preference shock remains the major factor with a long term contribution of 51 percent. However, the unconditional variance is again dominated by the cost push shock. The variance of the output gap is initially mainly due to the monetary policy shock and the technology shock, their contributions being 62 and 28 percent, respectively. The preference shock plays almost no role except for an initial five percent contribution that quickly fades out at longer horizons, as also do the two aforementioned shock, leaving room for the cost–push shock to become more and more important and resulting in a 92 percent contribution in the long term.

Table 3.3 presents the corresponding results for the model under adaptive learning. Interestingly, the preference and policy shocks play only a minor role for all variables. The variance of output growth is initially due to the technology shock with 53 percent and the cost–push with 41 percent shock. The impact of the latter increases to 85 percent in the long run reserving approximately the rest of the variance to the technology shock. The variances of inflation and the interest rate are dominated at all horizons by the cost–push shock. While both the technology and the policy shock exert almost no effect on the overall variance, the preference shock accounts for 3 and 8 percent at long horizons, respectively. The output gap is also dominated by the cost–push shock, although at short horizons, for $k = 1, 4$, the technology shock accounts for about 11–12 percent and the policy shock for about 9–10 percent.

3.6. Conclusions

This paper developed a general method to estimate DSGE models under the assumption of either rational expectations or alternatively adaptive learning. The latter explicitly takes account of the accumulation of information that gives agents a better understanding of the economy over time. Rational expectations can be interpreted as a special case of this imperfect knowledge approach where agents at each point knew the full sample estimates and used them for their forecasting procedures. However, adaptive learning does not offer additional degrees of freedom, but is a device that models agents perceptions. In the standard case the perceptions are assumed to coincide with those implied by the full sample model whereas the general AL case keeps tracks of agents' growing information set. The difference of both methods lies in the modelling of agents' forecasting model, which corresponds to the correct model given all information even at early stages where this information is not yet available in the RE case and the use of real–time models that change over time reflecting the growing knowledge over time in the AL case. Thus the RE method implicitly inserts the full sample forecast model into the expectations side of the model at any point in time whereas the AL method inserts the forecast model that an econometrician would recommend at a certain point in time given all available information. Since there are no more degrees of freedom, just a different sequence of forecast models, it is a priori unknown which of these approaches gives a better fit to the data. However, we find that the AL approach dominates the RE approach in terms of the maximized likelihood and conclude that the AL estimates offer a superior description of actual economic relations.

While some estimates qualitatively agree, we find that many key parameters are substantially different among the two different methods. A particularly important finding is that the backward looking term in the demand equation turns out to be very high implying that forward looking elements in the equation determining the output gap are not important, but instead that the gap is characterized by high structural persistence. This result clearly opposes the findings from the rational expectations model which come to the opposite conclusion.

As regards the assessment of the role of technology shocks, our results suggest that all shock terms are important, since all estimated standard deviations are highly significant. Quantitatively though, as demonstrated by the variance decompositions, the assessment of the importance of our four shocks differs for both methods.

In the RE model each shock has its role. While the cost–push shock is the dominant force except for the variance of output growth, to which the main contribution stems from preference shocks, policy shocks are important for the output gap, output growth and inflation, at least in the short term. Preference shocks are also important for variations in the interest rate. The effect of the technology shock, however, that contributes to output gap variability, to some extend also to output growth, and little to the inflation rate, is always dominated by at least one of the other shocks at any forecast horizon. Thus while technology shocks do have some explanatory power they are identified as the least important source of variation.

In the AL model the judgment is different. In this model preference shock exerts almost no effect at all, and to a lesser extend this holds also true for the monetary policy shock. Under adaptive learning, thus, we find that besides the important cost–push shock the only other influential source of volatility is the technology shock.

Quarters Ahead	Preference Shock	Cost–Push Shock	Technology Shock	Policy Shock
		Output Growth		
1	53.6	1.9	15.5	29.0
4	48.4	3.6	16.3	31.4
8	48.3	4.0	16.3	31.4
12	48.5	4.0	16.3	31.3
20	48.7	4.0	16.2	31.1
40	48.8	3.9	16.2	31.1
∞	48.8	3.9	16.2	31.1
		Inflation		
1	1.5	57.5	12.8	28.2
4	6.0	65.4	9.0	19.7
8	9.7	69.7	6.4	14.2
12	10.7	72.8	5.2	11.3
20	10.0	77.9	3.8	8.4
40	6.8	85.7	2.4	5.2
∞	0.0	100.0	0.0	0.0
		Interest Rate		
1	84.7	3.2	3.8	8.3
4	84.3	9.0	2.1	4.6
8	79.9	15.7	1.4	3.0
12	75.3	21.2	1.1	2.4
20	67.0	30.3	0.8	1.9
40	51.2	46.9	0.6	1.4
∞	0.3	99.7	0.0	0.0
		Output Gap		
1	5.1	4.1	28.4	62.4
4	2.8	32.5	20.2	44.5
8	1.8	60.4	11.8	26.0
12	1.4	72.3	8.0	17.6
20	0.9	83.6	4.8	10.6
40	0.5	91.7	2.4	5.3
∞	0.5	91.7	2.4	5.3

Table 3.2: Variance decomposition for the RE results.

Quarters Ahead	Preference Shock	Cost–Push Shock	Technology Shock	Policy Shock
		Output Growth		
1	2.2	40.5	52.5	4.9
4	1.1	74.7	21.9	2.4
8	0.7	84.6	13.1	1.6
12	0.7	84.9	12.8	1.6
20	0.7	84.9	12.8	1.6
40	0.7	84.9	12.8	1.6
∞	0.7	84.9	12.8	1.6
		Inflation		
1	0.7	98.1	0.7	0.6
4	1.0	97.5	0.8	0.7
8	1.2	97.2	0.9	0.7
12	1.4	97.1	0.9	0.7
20	1.8	96.8	0.8	0.7
40	2.5	96.1	0.8	0.6
∞	0.1	99.9	0.0	0.0
		Interest Rate		
1	1.0	98.1	0.5	0.4
4	1.5	97.9	0.3	0.2
8	2.5	96.8	0.4	0.3
12	3.3	95.9	0.4	0.3
20	4.7	94.6	0.4	0.3
40	7.6	91.7	0.4	0.3
∞	0.3	99.7	0.0	0.0
		Output Gap		
1	3.8	75.7	11.3	9.1
4	4.3	73.2	12.4	10.1
8	0.4	97.4	1.2	1.0
12	0.2	99.0	0.4	0.4
20	0.1	99.6	0.2	0.2
40	0.0	99.8	0.1	0.1
∞	0.0	99.8	0.1	0.1

Table 3.3: Variance decomposition for the AL results.

Concluding Remarks

The goal of this thesis was to investigate the empirical relevance of the adaptive learning approach. This method has experienced a recent boom culminating in a line of research that analyzes stability of systems under imperfect knowledge. The focus in on deriving conditions for policy that assure that additional dynamics resulting from the learning propagation mechanism fade out over time and imply convergence of economic systems to a perfect information, rational expectations equilibrium.

There is, however, little empirical treatment of this topic. Were the aforementioned dynamic propagation mechanism quantitatively negligible, the related convergence result would not appear to be relevant. We therefore assessed the role of adaptive learning in important empirical applications and tried to analyze the contribution of this approach to explaining phenomena in time series data.

Our first chapter showed how the assumption of adaptive learning endogenizes policy behavior, thus being able to explain changes in the stance of monetary policy that occurred e.g. in the US during the early 1980s under the tenure of Paul Volcker. These changes are usually traced back to exogenous forces such as the change in chairmanship of the Fed. The standard view that policy has not been conducted optimally must be modified. Indeed, we saw that taking account of adaptive learning completely explains the different stance in policy with inflation becoming a primary concern of central bankers in an environment under imperfect knowledge, since the beliefs of key parameters change substantially and imply a revision of the way policy is conducted. Our findings thus suggest that any policymaker would by the early 1980s have learned to behave in the way we observed in the history of US monetary policy, while conversely in the seventies any policymaker would have conducted policy in a way that would nowadays be criticized for having been non–optimal. This is an interesting result since this implies an equivalent change in policy for both the US and the UK economy, a feature that the change in chairmanship alone cannot address.

Our analysis revealed that historic interest rate decisions are highly correlated with optimal rates, but that even a policymaker under imperfect information would have reacted stronger towards the rise in inflation in the 1970s. On the other hand, policymaking in the last 25 years is consistent with an optimizing but learning policymaker. Policy reactions in the US and particularly in the UK until the early 1980s appear conservative related to recommendations of an optimizing decision maker. However, as the uncertainty surrounding the estimates of the model is initially high and substantially declines at the same time, we conjecture that accounting for uncertainty in the sense of Brainard is a key factor in policy setting.

We pursued this idea in the second chapter where the main focus is on the inflation–unemployment performance in an economy with an adaptively learning policymaker. Despite him behaving optimally the fact that his information set is limited almost inevitably leads the economy into great inflation like episodes. We find all characteristics that are observed in US and UK time series to be generated by model simulations under learning. These include the substantial revisions in perceived inflation persistence and the strong underestimation of the natural rate of unemployment in the late 1960s and 1970s. Furthermore, a Brainard–type policymaker will acknowledge the high uncertainty of econometric estimates until the early 1980s and if he acts in a conservative manner as advocated by Brainard, we consistently recover paths of inflation and unemployment as experienced in the US and the UK. We conclude that the adaptive learning hypothesis is capable of explaining important historic facts that are otherwise difficult to account for, and that in the empirical analysis uncertainty plays a central role.

Given the empirical relevance of the concept as documented in the two previous chapters, the third chapter developed a method that allows to estimate dynamic stochastic general equilibrium models under adaptive learning, replacing the assumption of rational expectations. We applied this approach to a version of the New Keynesian model. We found firstly, that the estimation results under the assumption of adaptive learning give a better fit to the data then under that of rational expectations, and secondly, that key results are substantially different under both methods. We concluded that the strong assumption of perfect knowledge which is implicit in the rational expectations framework introduces a bias into the estimation results.

To summarize, the adaptive learning approach has proven to be a powerful and empirically relevant tool, implementing an informationally plausible alternative, or extension, to the rational expectations hypothesis. We investigated the effect of the adaptive learning assumptions in important models of the economy and discussed how the standard assumption of rational expectations can be modified to account for the role of learning related dynamics. Models estimated under imperfect knowledge give a better fit to the data and reveal important insights otherwise masked by the rigid full information setup. We found robust evidence for the importance of adaptive learning with a particular emphasis on parameter uncertainty, the latter having not yet been taken account of in the related literature. It also turned out that estimation procedures involving the extraction of unobservable components emphasize a substantial role of misperceptions in these variables. Importantly, unlike knowledge on structural parameters, latent variables will always remain a source of uncertainty which can lead to suboptimal outcomes over and over again.

Although we investigated the impact of imperfect knowledge on optimal policy and on forecasting procedures in isolation, our results suggest that both channels are important and thus should be jointly analyzed. Thus, empirical analysis of models under adaptive learning should proceed by modelling learning as a two–sided process giving an important role to optimal policy decisions and appropriate forecast models under imperfect knowledge. Furthermore, the role of uncertainty of beliefs should be taken into consideration, since moderate behavior played an important role in the outbreak of the great inflation in the US and the UK.

Appendix

A.1. Recursive least squares and constant gain algorithms

We will discuss the relation of ordinary least squares, its recursive representation and constant gain algorithms (CGA) which are popular in the adaptive learning literature. As discussed in the text, constant gain algorithms are obtained from recursive least squares by replacing the decreasing gain factor by an appropriate constant; the literature has reached a consensus on the relevant range for this parameter and agrees on particular values. However, CGAs are recursive representations of the discounted least squares formula. Since the latter rarely appears in related work, we recall its definition

Definition 1. *The discounted least squares estimator for discount factor $\delta \leq 1$ of a vector of t observations y on the $t \times k$ regressor matrix x is given by the ordinary least squares estimator of the regression of Δy on Δx, where $\Delta = \operatorname{diag}(\delta^{t-1}, \delta^{t-2}, \ldots, \delta^1, \delta^0)$ is a diagonal matrix with geometrically increasing entries.*

Lemma 1. *Versions of the ordinary least squares estimator in the regression of the t–dimensional endogenous variable y on x are given by*

$$\beta_t = (X'\Delta^2 X)^{-1} X'\Delta^2 y = \big(\sum_{i=1}^{t} \delta^{2(t-i)} x_i x_i'\big)^{-1} \big(\sum_{i=1}^{t} \delta^{2(t-i)} x_i y_i\big) \tag{A.1}$$

The following equations are the corresponding recursive representation,

$$\beta_t = \beta_{t-1} + \gamma_t \Sigma_t^{-1} x_t'(y_t - x_t \beta_t) \tag{A.2a}$$

$$\Sigma_t = \Sigma_{t-1} + \gamma_t (x_t x_t' - \Sigma_{t-1}). \tag{A.2b}$$

For $\delta = 1$ and $\gamma = 1/t$, these equations yield the ordinary least squares estimator and its recursive version. If we choose an arbitrary δ, equations (A.1) *become discounted least squares, and an equivalent representation is given by formulae* (A.2) *if we choose* $\gamma = 1 - \delta^2$.

The least squares formula (A.1) is of course well known and is stated here only for completeness. Similarly, the discounted least squares version is a direct consequence of the definition. Only the equivalence of the recursive formulations to their non–recursive counterparts remains to be shown, in particular the relation of the discount factor and the gain parameter.

We are interested in terms of the form $c_t = R_t^{-1} S_t$, where both R_t and S_t are scaled sums, as in equation (A.1), and thus allow for recursive representations. To be concrete, we write $R_t = \gamma\alpha_t \sum_{i=1}^{t} g_i$ and $S_t = \gamma\alpha_t \sum_{i=1}^{t} h_i$. We focus on R_t first. After isolating the last summation term and factoring out α_t/α_{t-1} in the remaining sum, we can substitute in R_{t-1} to obtain the recursion $R_t = \alpha_t/\alpha_{t-1} R_{t-1} + \gamma\alpha_t g_t$.[ii]

Under either condition

$$\begin{aligned} \text{(i)} \quad & 1 - \alpha_t/\alpha_{t-1} = \gamma\alpha_t \\ \text{or (ii)} \quad & 1 - \alpha_t/\alpha_{t-1} = \gamma, \end{aligned}$$

we can deduce one of the more common terms for $R_t - R_{t-1}$ (by subtracting R_{t-1} from both sides of our recursion), namely $\gamma\alpha_t(g_t - R_{t-1})$ and $\gamma(\alpha_t g_t - R_{t-1})$, respectively.

Using similar steps and substituting in the corresponding formula for c_{t-1}, we can rewrite $S_t = \alpha_t/\alpha_{t-1} R_{t-1} c_{t-1} + \gamma\alpha_t h_t$. Collecting results and using footnote (ii), we finally arrive at the familiar updating term $c_t - c_{t-1} = \gamma\alpha_t R_t^{-1}(h_t - g_t c_{t-1})$.

To obtain versions of the least squares formula, we set $g_i = \delta^{-2i} x_i x_i'$, $h_i = \delta^{-2i} x_i y_i$. The choice $\delta = 1$, $\alpha_t = 1/t$, and $\gamma = 1$ yields ordinary least squares, that is $R_t^{-1} S_t$ corresponds to equation (A.1). Since these coefficients satisfy the first condition, we obtain the recursive least squares formulae, i.e. equations (A.2) with gain factor $1/t$. For arbitrary positive $\delta < 1$, the choice $\alpha_t = \delta^{2t}$ and $\gamma = 1 - \delta^2$, yields discounted least squares with discount factor δ, i.e. equation (A.1) for arbitrary discount factor. Since these parameters satisfy

[ii] This equation implies $\alpha_t/\alpha_{t-1} R_t^{-1} R_{t-1} = \mathrm{I} - \gamma\alpha_t R_t^{-1} g_t$.

the second condition, we obtain the constant gain algorithm with gain $1 - \delta^2$ as recursive representation, as in equations (A.2).

A.2. State space form of the dynamic program

We set

$$y_t = (\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, u_t, u_{t-1}, 1, u_t^*, u_{t-1}^*, V_{t-1})' \tag{A.3}$$

the dynamic program is given by

$$\max - \sum_{t \geq s} \delta^{t-s} [y_t' \Omega_{yy} y_t + u_t' \Omega_{uu} u_t + 2 y_t' \Omega_{yu} u_t] \tag{A.4}$$

$$\text{s.t.} \tag{A.5}$$

$$y_t = A y_{t-1} + B u_t \tag{A.6}$$

where

$$A = \left(\begin{array}{ccc|ccc|cc|cc} \alpha_1 & \alpha_2 & 0 & 0 & \theta_1 & \theta_2 & c_\pi & -\theta_1 & -\theta_2 & 0 \\ \hline \multicolumn{3}{c|}{I_3} & \multicolumn{3}{c|}{0_{3\times 3}} & \multicolumn{2}{c|}{0_{3\times 2}} & \multicolumn{2}{c}{0_{3\times 2}} \\ \hline -\lambda/4 & -\lambda/4 & -\lambda/4 & -\lambda/4 & \rho_1 & \rho_2 & c_u & 1-\rho_1 & -\rho_2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline \multicolumn{3}{c|}{0_{2\times 3}} & \multicolumn{3}{c|}{0_{2\times 3}} & \multicolumn{2}{c|}{I_2} & \multicolumn{2}{c}{0_{2\times 2}} \\ \hline \multicolumn{3}{c|}{0_{2\times 3}} & \multicolumn{3}{c|}{0_{2\times 3}} & \multicolumn{2}{c|}{\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}} & \multicolumn{2}{c}{0_{2\times 2}} \end{array}\right), \tag{A.7}$$

$$B = (0, 0, 0, 0, \lambda, 0, 0, 0, 0, 1)', \tag{A.8}$$

and

$$\Omega = \begin{pmatrix} \Omega_{yy} & \Omega yu \\ \Omega yu' & \Omega_{uu} \end{pmatrix} \tag{A.9}$$

consisting of the block matrices

$$\Omega_{yy} = \left(\begin{array}{c|c|c|c|c} \begin{matrix} 1 & -\pi^* \\ -\pi^* & 1 \end{matrix} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} \omega_x & -\kappa\omega_x \\ -\kappa\omega_x & \kappa^2\omega_x \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \omega_u \end{array}\right) \tag{A.10}$$

$$\Omega_{uu} = \omega_u, \tag{A.11}$$

and

$$\Omega_{yu} = (0, 0, 0, 0, 0, 0, 0, 0, 0, -\omega_{uu})' \tag{A.12}$$

A.3. Solution strategy

Our problem is to choose a sequence $\{v_t\}_{t\geq 0}$ to minimize

$$\sum_t \beta^t [x_t' R x_t + 2x_t' W u_t + u_t' Q u_t] \tag{A.13}$$

subject to the linear transition law

$$x_{t+1} = Ax_t + Bu_t + \varepsilon_{t+1} \tag{A.14}$$

where we adopt the notation of Hansen and Sargent (2004) to facilitate comparison. First, we note the well known result that the solution is given by $u_t = -Fx_t$, which is also established in the aforementioned monograph, with F given by

$$F = (Q + \beta B' P B)^{-1} (\beta B' P A + W') \tag{A.15}$$

Solving the policymaker's problem thus reduced to computing the matrix P, which is known to satisfy the algebraic Riccati equation

$$P = R + \beta A' P A - (\beta A' P B + W)(Q + \beta B' P B)^{-1} (\beta B' P A + W') \tag{A.16}$$

We will apply an efficient routine provided by Evan W. Anderson.[iii] This method is designed for undiscounted optimal linear regulator problems without mixed terms, so several transformations as suggested by Hansen and Sargent are in order. We set

$$A^* = \sqrt{\beta}(A - BQ^{-1}W') \tag{A.17a}$$

$$B^* = \sqrt{\beta} B \tag{A.17b}$$

$$R^* = R - WQ^{-1}W' \tag{A.17c}$$

$$Q^* = Q \tag{A.17d}$$

[iii] The matlab file schurgaux is available on his website http://www.math.niu.edu/~anderson.

and verify by inserting the matrices that our above system is equivalent to the undiscounted problem without mixed terms under certainty,

$$\sum_t [\hat{x}_t' R^* \hat{x}_t + \hat{v}_t' Q^* \hat{v}_t] \tag{A.18}$$

subject to

$$\hat{x}_{t+1} = A^* \hat{x}_t + B^* \hat{v}_t \tag{A.19}$$

This system gives rise to a corresponding Riccati equation in P^*,

$$P^* = R^* + A^{*\prime} P^* A^* - A^{*\prime} P^* B^* (Q^* + B^{*\prime} P^* B^*)^{-1} B^{*\prime} P^* A^*. \tag{A.20}$$

Conveniently, the solution to the latter Riccati equation coincides with the one in our original problem, a fact we capture in

Lemma 2. *The Riccati matrix of our original problem (xx) coincides with the Riccati matrix of the transformed system, $P = P^*$.*

Hansen and Sargent (2004) and Hansen and Sargent (2005) show how the problem (A.18) and (A.19), and thus implicitly equation (A.20), can be solved using a deflating subspace method. This involves computing the ordered generalized real Schur decomposition of the matrix pencil $\lambda L - N$, $\lambda \in \mathbb{C}$ with

$$L = \begin{pmatrix} I_n & B^* Q^{*-1} B^{*\prime} \\ 0_n & A^{*\prime} \end{pmatrix}, \quad N = \begin{pmatrix} A^* & 0_n \\ -R^* & I_n \end{pmatrix}. \tag{A.21}$$

Anderson's routine is designed to compute this decomposition, associated with our transformed problem. It yields a matrix, which after partitioning into 2×2 equally sized submatrices U_{ij}, $i, j = 1, 2$, let's us deduce the Riccati matrix $P = U_{21} U_{11}^{-1}$, and thus from lemma 2 the Riccati matrix of our original problem, which is then subsequently used to infer the optimal policy from equation (A.15).

It just remains us to provide the proof of our lemma, which is done by substituting the transformed system's matrices (A.17) into the original Riccati equation (A.16) and investigating each of the three terms on the right hand side of the above equation separately.

The first only requires substitution of R,

$$R = R^* + W Q^{-1} W', \tag{A.22}$$

the second term becomes the second term becomes

$$\begin{aligned}
&\beta A'PA \\
&=\beta(\beta^{-1/2}(A^* + B^*Q^{-1}W'))'P(\beta^{-1/2}(A^* + B^*Q^{-1}W')) \\
&=[A^{*\prime} + WQ^{-1}B^{*\prime}]P[A^* + B^*Q^{-1}W'] \\
&=A^{*\prime}PA^* + A^{*\prime}PB^*Q^{-1}W' + WQ^{-1}B^{*\prime}PA^* + WQ^{-1}B^{*\prime}PB^*Q^{-1}W'. \qquad \text{(A.23)}
\end{aligned}$$

Tedious computations are necessary for the third term. We start with the last two of its three factors, and use the shorthand notation $M \equiv (Q+B^{*\prime}PB^*)$:

$$\begin{aligned}
&M^{-1}[B^{*\prime}P(A^* + B^*Q^{-1}W') + W'] \\
&= M^{-1}[B^{*\prime}PA^* + (B^{*\prime}PB^* + Q)Q^{-1}W'] \\
&= M^{-1}[B^{*\prime}PA^* + MQ^{-1}W'] = M^{-1}B^{*\prime}PA^* + Q^{-1}W'
\end{aligned}$$

Hence

$$\begin{aligned}
&[(A^{*\prime} + WQ^{-1}B^{*\prime})PB^* + W] \times M^{-1}[B^{*\prime}P(A^* + B^*Q^{-1}W') + W'] \\
&= [A^{*\prime}PB^* + WQ^{-1}(Q + B^{*\prime}PB^*)] \times [M^{-1}B^{*\prime}PA^* + Q^{-1}W'] \\
&= A^{*\prime}PB^*M^{-1}B^{*\prime}PA^* + A^{*\prime}PB^*Q^{-1}W' + WQ^{-1}B^{*\prime}PA^* \\
&\quad + WQ^{-1}B^{*\prime}PB^*Q^{-1}W' + WQ^{-1}W' \qquad \text{(A.24)}
\end{aligned}$$

The second term of (A.24) cancels with the second of (A.23) and similarly the third and fourth terms cancel. The fifth term cancels with the second in (A.22). Collecting the remaining terms, which are the first expressions of each separate term, we obtain

$$P = R^* + A^{*\prime}PA^* - A^{*\prime}PB^*(Q + B^{*\prime}PB^*)^{-1}B^{*\prime}PA^* \qquad \text{(A.25)}$$

Hence, the Riccati matrix of the untransformed system also satisfies the Riccati equation of the transformed system (A.20), thus from the uniqueness of this matrix we deduce $P = P^*$ as was to be shown.

Bibliography

Andrews, Donald W. K. and Ray C. Fair (1988): "Inference in Nonlinear Econometric Models with Structural Change." *Review of Economic Studies*, 55:615–639.

Ball, Laurence and N. Gregory Mankiw (2002): "The NAIRU in Theory and Practice." *Working Paper*.

Barro, Robert J. and David B. Gordon (1983): "Rules, discretion and reputation in a model of monetary policy." *Journal of Monetary Economics*, 12:101–121.

Batini, Nicoletta and Jennifer V. Greenslade (2006): "Measuring the UK short-run NAIRU." *Oxford Economic Papers*, 58:28–49.

Blanchard, Olivier J. and Charles M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations." *Econometrica*, 48:1305–1311.

Blinder, Alan S. (1997): "Distinguished Lecture on Economics in Government: What Central Bankers Could Learn from Academics–And Vice Versa." *Journal of Economic Perspectives*, 11:3–19.

Brainard, William (1967): "Uncertainty and the effectiveness of policy." *American Economic Review*, 57:411–425.

Bullard, James and Stefano Eusepi (2005): "Did the Great Inflation Occur Despite Policymaker Commitment to a Taylor Rule?" *Review of Economic Dynamics*, 8:324–359.

Bullard, James and Kaushik Mitra (2002): "Learning about Monetary Policy Rules." *Journal of Monetary Economics*, 49:1105–1129.

Calvo, Guillermo A. (1983): "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics*, 12:382–398.

Cho, In-Koo, Noah Williams and Thomas J. Sargent (2002): "Escaping Nash Inflation." *Review of Economic Studies*, 69:1–40.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (1997): "Sticky price and limited participation models of money: A comparison." *European Economic Review*, 41:1201–1249.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (1999): "Monetary Policy Shocks: What have we learned and to what end?" In Taylor, John B. and Michael Woodford, editors, *Handbook of Macroeconomics, Volume 1*, pages 65–148. Elsevier Science B.V.

Christiano, Lawrence J. and Christopher J. Gust (1999): "Taylor Rules in a Limited Participation Model." *NBER Working Paper*, 7017.

Clarida, Richard, Jordi Galí and Mark Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Theory." *Quarterly Journal of Economics*, 115:147–180.

Cogley, Timothey and Thomas J. Sargent (2001): "Evolving Post-World War II U.S. Inflation Dynamics." *NBER Macroeconomic Annual*, pages 331–373.

Evans, George and Seppo Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Frontiers of Economic Research. Princeton University Press, Princeton.

Franz, Wolfgang (2005): "Will the (German) NAIRU Please Stand Up?" *German Economic Review*, 6:131–153.

Fuhrer, Jeff and George Moore (1995): "Inflation Persistence." *The Quarterly Journal of Economics*, 110:127–59.

Gerlach, Stefan and Lars E. O. Svensson (2003): "Money and Inflation in the Euro Area: A Case for Monetary Indicators?" *Journal of Monetary Economics*, 50:1649–1672.

Goodfriend, Marvin (1993): "Interest Rate Policy and the Inflation Scare Problem: 1979–1992." *FRB Richmond Economic Quarterly*, 79.

Gordon, Robert J. (1997): "The Time-Varying NAIRU and Its Implications for Economic Policy." *Journal of Economic Perspectives*, 11:11–32.

Greenslade, Jennifer V., Richard G. Pierse and Jumana Saleheen (2003): "A Kalman filter approach to estimating the UK NAIRU." *Bank of England Working Paper*.

Hamilton, James D. (1994): *Time Series Analysis*. Princeton University Press.

Hansen, Lars Peter and Thomas J. Sargent (2004): *Recursive Models of Dynamic Linear Economies*. Monograph.

Hansen, Lars Peter and Thomas J. Sargent (2005): *Misspecification in Recursive Macroeconomic Theory*. Monograph.

Harvey, Andrew C. (1993): *Time Series Models*. MIT Press, 2nd edition.

Harvey, Andrew C. (2003): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge Universtiy Press.

Hodrick, Robert J. and Edward C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking*, 29:1–16.

Ireland, Peter N. (2003): "Irrational expectations and econometric practice: discussion of Orphanides and Williams, 'Inflation scares and forecast-based monetary policy'." Working Paper 2003–22, Federal Reserve Bank of Atlanta.

Ireland, Peter N. (2004): "Technology Shocks in the New Keynesian Model." *The Review of Economics and Statistics*, 86:923–936.

Junker, Simon (2008): "Adaptive Learning in the Empirical Analysis." *Working Paper*.

Kalman, Rudolf E. (1960): "A new Approach to Linear Filtering and Predictive Problems." *Transactions ASME, Journal of basic engineering*, pages 34–45.

Klein, Paul (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model." *Journal of Economic Dynamics and Control*, 24:1405–1423.

Kreps, David (1998): "Anticipated Utility and Dynamic Choice." In Jacobs, Donald P., Ehud Kalai, Morton I. Kamien and Nancy L. Schwartz, editors, *Frontiers of Research in Economic Theory*, pages 242–274. Cambridge University Press.

Lucas, Robert E. Jr. (1972): "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4. 103–124.

Lucas, Robert E. Jr. (1976): "Econometric Policy Evaluation: A Critique." In *Carnegie Rochester Conference Series on Public Policy*. North–Holland.

Milani, Fabio (2005): "Adaptive learning and inflation persistence." *Working Paper*.

Muth, John F. (1961): "Rational Expectations and the Theory of Price Movements." *Econometrica*, 29:315–335.

Okun, Arthur (1962): "Potential Output: Its Measurement and Significance." *American Statistical Association 1962 Proceedings of the Business and Economic Section*.

Orphanides, Athanasios (2002): "Monetary Policy Rules and the Great Inflation." *FRS Working Paper*.

Orphanides, Athanasios and Simon van Norden (2002): "The Unreliability of Output–Gap Estimates in Real Time." *The Review of Economics and Statistics*, 84:569–583.

Orphanides, Athanasios and John C. Williams (2005): "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control*, 29:1927–1950.

Orphanides, Athanasios and John C. Williams (2007): "Robust monetary policy with imperfect knowledge." *Journal of Monetary Economics*, 54:1406–1435.

Primiceri, Giorgio (2006): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy." *The Quarterly Journal of Economics*, 121:867–901.

Rotemberg, Julio J. (1982): "Sticky Prices in the United States." *Journal of Political Economy*, 90:1187–1211.

Rudebusch, Glenn D. and Lars E. O. Svensson (1998): "Policy Rules for Inflation Targeting." NBER Working Papers 6512, National Bureau of Economic Research.

Sack, Brian (2000): "Does the fed act gradually? A VAR analysis." *Journal of Monetary Economics*, 46:229–256.

Schorfheide, Frank (2005): "Learning and Monetary Policy Shifts." *Review of Economic Dynamics*, 8:392–419.

Sims, Christopher A (2002): "Solving Linear Rational Expectations Models." *Computational Economics*, 20:1–20.

Smets, Frank (2002): "Output Gap Uncertainty: Does it matter for the Taylor Rule?" *Empirical Economics*, 27:113–129.

Taylor, John B. (1993): "Discretion versus policy rules in practice." *Carnegie–Rochester Conference Series on Public Policy*, 39:195–214.

Taylor, John B. (1999): "How the Rational Expectations Revolution Has Changed Macroeconomic Policy Research." *Lecture presented at the 12th*

World Congress of the International Economic Association.

Uhlig, Harald (1995): "A toolkit for analyzing nonlinear dynamic stochastic models easily." *Federal Reserve Bank of Minneapolis Discussion Paper 101*.

Woodford, Michael (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, New Jersey.