Cardiff Economics Working Papers

Cardiff Business School
Cardiff University

Ysgol Fusnes Caerdydd
Prifysgol Caerdydd

Working Paper No. E2013/14

Stabilization policy, rational expectations and price-level versus inflation targeting: a survey

Michael C. Hatcher and Patrick Minford

December 2013

Cardiff Business School
Aberconway Building
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Stabilization policy, rational expectations and price-level versus inflation targeting: a survey

Michael Hatcher*
University of Glasgow

Patrick Minford
Cardiff University and CEPR

26 Dec 2013

Abstract

We survey recent literature comparing inflation targeting (IT) and price-level targeting (PT) as macroeconomic stabilization policies. Our focus is on New Keynesian models and areas which have seen significant developments since Ambler's (2009) survey: the zero lower bound on nominal interest rates; financial frictions; and optimal monetary policy. Ambler's main conclusion that PT improves the inflation-output volatility trade-off in New Keynesian models is reasonably robust to these extensions, several of which are attempts to address issues raised by the recent financial crisis. The beneficial effects of PT therefore appear to hang on the joint assumption that agents are rational and the economy New Keynesian. Accordingly, we discuss recent experimental and survey evidence on whether expectations are rational, as well as the applied macro literature on the empirical performance of New Keynesian models. In addition, we discuss a more recent strand of applied literature that has formally tested New Keynesian models with rational expectations. Overall the evidence is not conclusive, but we note that New Keynesian models are able to match a number of dynamic features in the data and that behavioral models of the macroeconomy are outperformed by those with rational expectations in formal statistical tests. Accordingly, we argue that policymakers should continue to pay attention to PT.

1 Introduction

The past two decades have seen a resurgence of interest in optimal monetary policy, including research on potential alternatives to inflation targeting (IT). In this survey, we focus on an alternative policy that is considered a serious contender to IT: price-level targeting (PT). Under a PT regime, the central bank attempts to stabilize the aggregate price level (eg the CPI) around a predetermined target price path. Hence, for example, if there is an inflationary shock

*We would like to thank Tony Yates and Alfred Duncan for useful comments.

that takes the price level above the target price path, below-average inflation is required next period in order to return the price level to target. This contrasts with an IT regime where inflation stabilization is the goal of policy and above average inflation today would be followed by average (ie on-target) inflation tomorrow. In other words, 'bygones are bygones' under IT, but the opposite is true under PT.

Our survey begins by discussing recent developments in the PT literature on macro stabilization. We focus on New Keynesian models and important developments since Ambler's (2009) survey, which can be grouped under three headings: optimal monetary policy; the zero lower bound on nominal interest rates; and financial frictions. Several of these recent developments are initial attempts to address issues raised by the financial crisis. The main theoretical finding of the survey is that Ambler's original conclusion that PT can improve the inflation-output volatility trade-off in New Keynesian models remains intact, albeit that the list of caveats is much longer. In particular, we note an interesting split in the literature: PT consistently outperforms IT in New Keynesian models where policymakers commit to instrument rules of the kind advocated by Taylor (1993), but results in favour of PT which are derived from optimal policies – under either commitment or discretion – are generally quite fragile to the assumptions one makes about the economic environment faced by the policymaker. We also argue, however, that the second case is the one of least interest for real-world central banks, because optimal policies do not perform well across a range of alternative models of the economy and are more difficult to implement than simple Taylor-type rules. In this sense the performance of PT in New Keynesian models is fairly robust to extensions to the basic framework. This encourages us to focus on the key assumptions on which the beneficial effects of PT are known to hang – are economic agents rational and the economy New Keynesian?

With this question in mind, the second half of the survey considers recent experimental and survey literature on rational expectations, and the applied macro literature on the empirical performance of New Keynesian models. Both these areas have seen important developments in recent years and they offer different insights. On the one hand, experiments and surveys tend to reject the strict rational expectations hypothesis, but we note that there are potential problems with micro evidence of this kind, several of which we discuss in detail. On the other hand, the recent literature on the empirical performance of New Keynesian models initiated by Christiano et al. (2005) is broadly supportive but should be treated with caution because it does not amount to a formal statistical test of New Keynesian models with rational expectations. There is, however, a more recent strand of applied macro literature that has formally tested New Keynesian models and alternative expectational assumptions. We explain the intuition and methodology behind this approach before discussing some of the main results from this growing literature. Overall, the evidence on rational expectations and New Keynesian models is far from clear-cut, but we note that these models are able to match a number of important dynamic features in the data, and that behavioral models of the macroeconomy seem to be outperformed by those with rational expectations in formal statistical tests. We also argue that when considering macroeconomic policy and its effects on the economy, priority should be given to macroeconomic evidence in cases where it is able to establish results on strong statistical grounds.

In a concluding section we consider the policy implications of the survey. We think there are two main lessons. First, from a theoretical perspective, the performance of PT is surprisingly robust to initial attempts to address issues raised by the financial crisis by extending the standard New Keynesian model. If anything, then, the events of the crisis appear to have strengthened the case for PT – at least based on the available theoretical models. As Walsh (2011) cautions, this literature is at an early stage, with many of the key trade-offs raised by the crisis not yet fully understood by policymakers and academics. Building models that can speak to these issues should therefore be high on the research agenda and, in time, these models should become useful tools to (re-)assess the case for a change in monetary policy regime to PT. The second lesson is that there is a strong disconnect between experimental and survey evidence on rational expectations on the one hand, and the applied macro literature on the empirical performance of New Keynesian models on the other. However, since behavioral models of the macroeconomy are outperformed by those with rational expectations in formal statistical tests at the macro level, and there are serious potential problems with surveys and experiments at the micro level, we argue that policymakers should continue to pay attention to PT. Indeed, the available applied macro evidence suggests that widespread concerns about rational expectations and New Keynesian models in the aftermath of the financial crisis may be unwarranted.

2 The baseline New Keynesian model

2.1 The model

The distinguishing feature of New Keynesian models is that monopolistically-competitive firms set their prices optimally to maximise profits, subject to constraints on how frequently they can re-set prices or how costly it is to do so. In the baseline specification due to Calvo (1983), each firm can change its price with a constant probability, so that the interval between price changes is a random variable. Profit maximization by firms leads to a first-order condition that relates current price to a mark-up on marginal cost and the expected future price. Log-linearizing this first-order condition around a zero-inflation steady-state and aggregating across firms leads to the New Keynesian Phillips curve in which economy-wide inflation, denoted π_t, is related to the output gap and expected future inflation:

$$\pi_t= \beta E_t\pi_{t+1}+\kappa x_t+u_t \tag{1}$$

where x_t is the output gap, β is the discount factor of the representative household (the sole owner of firms), $\kappa > 0$ is the slope of the Phillips curve, and u_t is a cost-push shock that follows an AR(1) process.

Equation (1) states that current inflation depends positively on (rationally) expected future inflation and the current output gap. So the main way in which the New Keynesian Phillips curve differs from a more traditional neoclassical Phillips curve is that inflation depends on expected future inflation (ie $E_t\pi_{t+1}$), rather than expected current inflation ($E_{t-1}\pi_t$).[1] This difference has important implications for monetary policy, since it implies that current inflation depends on the expected present discounted value of future output gaps and cost-push shocks, so that managing expectations about the future is important for current inflation control. This feature of the model is empirically plausible because there is substantial evidence that anticipated changes in monetary policy have real effects (eg Christiano et al. 2005).

The demand side of the model is standard: consumers maximise utility by choosing asset holdings optimally, giving rise to a Euler equation for each household. If we aggregate across households we get a single consumption Euler equation that describes the dynamics of aggregate consumption as a function of the real interest rate and expected future consumption. Log-linearizing this equation and imposing the market-clearing condition that consumption equals output minus government spending, we get the IS curve in the baseline New Keynesian model:

$$x_t = E_t x_{t+1} - \sigma(R_t - E_t\pi_{t+1}) + g_t \tag{2}$$

where $\sigma > 0$ is the intertemporal elasticity of substitution, x_t is the output gap, R_t is the nominal interest rate, and g_t is a government spending shock that follows an AR(1) process.

Like the New Keynesian Phillips curve, the IS curve is forward-looking: it emphasizes the importance of expectations about the future in the determination of current outcomes. In particular, the current output gap rises with the expected future output gap due to consumers' desire to smooth consumption, and an increase in expected future inflation raises the current output gap because this implies a fall in the real rate of interest (for any given value of the nominal interest rate).

In addition to the microfounded supply side, an important feature of the New Keynesian model is that a social loss function in inflation and output gap variations can be derived as an approximation to the utility function of the representative household; see, for example, Walsh (2010, Ch. 8.6). Consequently, social welfare analyses in these models are consistent with the model itself as

[1] We follow the convention that E_t is the mathematical expectations operator conditional on information available up to and including period t.

well as being broadly consistent with the stated objectives of real-world central banks. The approximate social loss function in the baseline model can be represented as follows:

$$L_t = E_t \sum_{j=0}^{\infty} \beta^j (\pi_{t+j}^2 + \alpha x_{t+j}^2) \tag{3}$$

where $\alpha > 0$ is the relative weight on output gap variations, which depends upon the parameters of the model.

This social loss function has two important roles in the literature. Firstly, by minimizing Equation (3) subject to (1) and (2) we can derive the first-order conditions under optimal policy. These first-order conditions provide insight into the main features of effective monetary policymaking and allow us to analyze the optimal response of policy to shocks that hit economy (see eg Walsh 2010, Ch. 8.4.3). Second, Equation (3) can be used as a basis for evaluating and comparing social welfare across alternative monetary policy regimes like IT and PT as, for example, in Vestin (2006).

The model is closed by introducing an additional equation that pins down the nominal interest rate R_t. This can be done in two ways. The first way, which has been popular in applied literature, is to specify an explicit instrument rule for the nominal interest rate. The most commonly used rule of this kind is the Taylor rule (Taylor, 1993):

$$R_t = c_\pi \pi_t + c_x x_t \tag{4}$$

where c_π, $c_x > 0$ are the reaction coefficients on inflation and the output gap, respectively; we treat the inflation target and steady state interest rate as zero here and in what follows.

Taylor rules are generally suboptimal – ie they do not minimize Equation (3). Nevertheless, they have become a popular way of modelling monetary policy, because they are easy to communicate; more robust across alternative models than optimal policies; and a useful way of describing the behaviour of central banks in estimated dynamic general equilibrium models.

The second way of closing the model is with an implicit instrument rule. To implement this approach a particular relationship between the endogenous variables is assumed to hold and this relationship is then added to the model as the third equation. For each such feasible relationship there will be an implied interest rate rule that implements it, but these rules are typically difficult to derive analytically. The interest rate rule itself is therefore usually left unspecified (hence the label 'implicit'), and the path of the interest rate is instead solved for numerically.

This approach has been popular in the theoretical literature where researchers derive first-order optimality conditions (under commitment or discretion) and then study optimal policy by assuming that the central bank sets nominal interest rates at whatever level is consistent with that policy. For example, in the case of the baseline New Keynesian model under optimal commitment, Evans and Honkapohja (2006) show that the first-order optimality condition and the implied interest rate rule are as follows:

$$\pi_t = -\frac{\alpha}{\kappa}(x_t - x_{t-1}) \tag{5}$$

$$R_t = c_1 x_{t-1} + c_2 g_t + c_3 u_t \tag{6}$$

where c_1, c_2 and c_3 are complicated functions of the model parameters.

2.2 Calvo contracts and micro-evidence on sticky prices

There is considerable empirical evidence suggesting that prices are 'sticky' in nominal terms as suggested by the Calvo model. In a seminal paper, Bils and Klenow (2004) conclude that average price spells in the US last between 3 and 4 months, but subsequent results in Nakamura and Steinsson (2008) that exclude sales prices point to longer price spells of 7 to 11 months – a result which is more consistent with early empirical work on price stickiness in the US. Dhyne et al. (2006) find similar results for the Euro Area as part of research conducted by the Inflation Persistence Network at the ECB: the average duration of price spells is 10 to 13 months. Consistent with these findings, the Calvo reset probability is typically calibrated to imply average price spells of between 2 and 4 quarters.

The factors responsible for nominal price rigidities remain somewhat of a mystery, however. Under Calvo price-setting these factors are not modeled: firms cannot change prices in some periods, and the probability of being able to change price is independent of the date at which prices were last re-set. Clearly, these are not realistic assumptions: it must be the case that firms choose to change prices in some periods and not to do so in others; and firms who have maintained the same price level for several periods are more likely to change their prices (ie re-optimize) than those who re-set their price recently. There is however an important advantage of the simplfying assumptions under Calvo: they facilitate aggregation by making the model tractable.[2] It is also important to note that the Calvo specification implies that there is a distribution of price changes across firms as in the data (see eg Bils and Klenow, 2004) – that is, some firms change price more frequently than others. Therefore, the Calvo model captures some major features of real-world price-setting within an analytically tractable framework.

Calvo contracts have been criticized in other ways, however, such as the assumption of a constant reset probability, which cannot explain why prices

[2] A natural alternative to Calvo price-setting is the Taylor model (Taylor, 1980) of overlapping fixed-duration price contracts, but this specification is less convenient analytically because it requires as many groups of price-setters as price-spell durations.

are reset more frequently in high-inflation environments. Several alternative forms of nominal rigidity have been proposed in the literature, including state-dependent pricing models (eg Golosov and Lucas, 2007) whereby firms change prices more frequently when it is more profitable to do so; and generalized Taylor models (Dixon and Le Bihan, 2012) where there are several sectors with Taylor (1980) contracts of different durations.[3] These alternative models appear to match the data better in some respects. Nevertheless, the baseline New Keynesian model with Calvo price-setting assumptions remains the dominant paradigm in the literature, and it is also the one that has been tested most extensively.

We may also say the following in defence of the baseline New Keynesian model. First, although the Calvo model does not provide any theoretical justification for the stickiness of prices, it is not clear that such a justification is necessary if our aim to match the dynamics of aggregate macro data. That is to say, the crucial variable from a macro perspective might be the duration for which prices are fixed (with the reasons being largely irrelevant).[4] Second, even if one were to accept the argument that Calvo contracts do not provide an adequate approximation to firms' pricing behaviour, one is left asking what we should to do about this. The best approach, in the absence of an alternative workhorse model that is clearly superior, would presumbly be to proceed as the current literature has – stressing that our conclusions should be treated with caution while staying open to the possibility of refinements or alternatives that could fit the facts better.

Since there is no clear consensus on an alternative model at the current time, the argument for moving away from the baseline New Keynesian model with Calvo price-setting is fairly weak. In fact, as we discuss below, an extended version of the baseline New Keynesian model provides a suprisingly good fit to the dynamic behaviour of key macro variables in the postwar period – and in particular during the Great Moderation. Having introduced the baseline New Keynesian model and discussed its main strengths and weaknesses, we now turn to literature that has used these models to assess the relative merits of PT from a stabilization perspective.

3 Macro stabilization literature

As Ambler (2009) points out, PT was initially motivated as a way of providing the economy with long-term price stability. Nevertheless, most of the recent interest in PT comes from its implications for short run macro stabilization, that

[3] For a useful discussion of state-dependent pricing models, see Walsh (2010, Ch. 6.2).

[4] Caplin and Spulber (1987) build a state-dependent pricing model where there is price stickiness at the firm level but none in the aggregate price level, so that the reasons for price stickiness are crucial for its aggregate implications. A tractable model of this kind is set out Lucas and Golosov (2007). However, Costain and Nakov (2011) develop a generalized model that nests the Calvo model and the Golosov-Lucas model and find that the best fit to micro pricing data arises when the degree of state dependence is low.

is, economic stability at business cycle frequencies. In fact, the recent resurgence in interest in PT has gone hand-in-hand with the rise in popularity of New Keynesian models, since there are welfare gains from 'history dependence' in a forward-looking, rational expectations environment (Woodford, 2003). History dependence means that monetary policy should respond systematically to past economic conditions, as well current and expected future economic outcomes. PT is an example of a history dependent policy because the central bank is obliged to offset past deviations from its target price path. As a result, PT will tend to produce superior stabilization outcomes to policies such as IT that lack history dependence.

In this section we discuss the main findings from the macro stabilization literature on IT versus PT, as surveyed by Ambler (2009). We then turn to more recent contributions that shed light on robustness to extensions of the baseline New Keynesian model. To keep contact with the rapidly-expanding literature in this area we split the discussion that follows into three subsections: optimal policy; the zero lower bound on nominal interest rates; and financial frictions.

3.1 Optimal policy

In a seminal paper, Kydland and Prescott (1977) showed that the distinction between discretion and commitment is crucial in rational expectations models. Under discretion, the policymaker optimizes on a period-by-period basis. As a result, it cannot make binding promises about the future and must take future expectations as given. This inability to commit generally leads to suboptimal outcomes, because policy cannot influence future expectations and so effectively has fewer instruments available to achieve its objectives. A classic example of a suboptimal outcome under discretion is the 'inflation bias' problem highlighted by Barro and Gordon (1983), whereby an equilibrium with steady-state inflation above the socially-optimal level results from a desire to push output above the natural rate. An efficient equilibrium outcome can be restored in various ways, including delegation to a central bank that is more inflation-averse than society (Rogoff, 1985); one which is subject to punishment via an optimal contract if the inflation target is not met (Walsh, 1995); or delegation to central bank with a PT mandate (Svensson, 1999).

It is important to note that discretion leads to suboptimal outcomes in dynamic models with rational expectations, even if the central bank *does not* have an over-ambitious output target (the assumption made in almost all of the recent literature, and in all the papers we review here). Svensson (1997) and Clarida et al. (1999) call this phenomenon 'stabilization bias' because it implies that the central bank's response to shocks that hit the economy is suboptimal, so that inflation and the output gap deviate from their first-best outturns. As with the inflation bias, stabilization bias can be eliminated if the central bank can make a binding commitment over the future path of monetary policy. Stabilization bias also implies that delegating monetary policy to a central bank without IT preferences could potentially improve social outcomes.

3.1.1 Optimal policy under discretion

Svensson (1999) was the first to formally investigate whether delegating monetary policy to a central bank with a PT mandate could improve stabilization outcomes under rational expectations. Using a model with a Lucas-type (or Neoclassical) Phillips curve, he showed that delegating PT preferences to the central bank delivers a 'free lunch' result: inflation variability is lower for any given level of output gap variability if output is sufficiently persistent.[5] In effect, PT eliminates some of the useless discretion present under IT because the central bank is required to undo deviations from its target price path in the future, so that discretionary behaviour becomes more costly and is curtailed. The crucial distinction is that 'bygones are bygones' under IT, whereas PT makes the objective function of a discretionary policymaker depend in part upon past outcomes, so that discretion carries a penalty which is absent under IT. Although Svensson's model features rational expectations, it is subject to two important limitations. First, the economy is described by a Phillips curve which performs poorly from an empirical perspective because it implies that only unanticipated changes in policy have real effects. Second, the social loss function Svensson uses is ad hoc in the sense that it cannot be derived from the utility function of a representative agent. As noted in Section 2.1, the New Keynesian model can address both these criticisms.

Vestin (2006) assesses the performance of PT in the baseline New Keynesian model. He shows that the free lunch result remains intact. Intuitively, since firms in New Keynesian models set current prices as a function of expected future prices, the extent to which current prices rise in response to an inflationary shock depends upon the impact of policy upon expectations about future inflation. Under IT, inflation expectations are effectively fixed on the inflation target because 'bygones are bygones'. Under PT, by contrast, firms expect a rise in current prices to be followed by a contraction in demand and lower future prices, in order to return the price level to its target path. As a result, firms raise their prices less in response to inflationary (ie cost-push) shocks under a PT regime, implying lower inflation variability for any given level of output gap variability.

The free lunch result is also robust to indexation to past inflation in periods when prices-setters are unable to reoptimize, except in the special case of full indexation (Röisland, 2006; Gaspar et al. 2007). Indexing prices to past inflation gives rise to the so-called 'hybrid New Keynesian Phillips curve' where inflation is additionally related to lagged inflation:

$$\pi_t - \gamma\pi_{t-1} = \beta E_t(\pi_{t+1} - \gamma\pi_t) + \kappa x_t + u_t \tag{7}$$

[5] PT dominates IT if the degree of endogenous persistence in the output gap exceeds 1/2.

where $0 \leq \gamma \leq 1$ represents the degree of indexation to past inflation, and $\pi_t - \gamma\pi_{t-1}$ is the 'quasi-difference' in inflation.

Although this specification performs better empirically than the standard New Keynesian Phillips curve, it is difficult to justify from a theoretical perspective because price-setters index in a purely backward-looking manner. Moreover, Steinsson (2003) shows that one can also justify a partially backward-looking New Keynesian Phillips curve by assuming that a fraction of price-setters are non-rational and follow a simple rule-of-thumb when setting prices, but in this case optimal policy implies base-level drift in the price level. Nevertheless, the numerical results in Nessén and Vestin (2005) suggest that PT will continue to dominate IT in welfare terms as long as the fraction of rule-of-thumb price-setters in the population is smaller than the fraction of rational price-setters.[6]

Blake et al. (2011) tackle a somewhat different issue that arises under discretion. Since there are multiple rational expectations equilibria under discretion when the baseline New Keynesian model is augmented with capital accumulation (Blake and Kirsanova, 2012), they demonstrate that IT can potentially outperform PT under discretion in these models, because the lowest welfare equilibrium under delgation to a PT central bank may be inferior to the highest welfare equilbrium attainable under IT. Consequently, it is generally not possible to predict whether PT will outperform IT in welfare terms – at least in the absence of a mechanism for selecting one equilibrium over another.[7] Finally, Masson and Shukayev (2012) show that if the central bank operates under discretion and the public believes that there is possibility it will rebase the price-level target in response to large shocks, there can be multiple equilbria under PT even in the absence of capital accumulation. As a result, the economy could end up in a low credibility PT equilibrium (ie one with a high probability of rebasing) where output volatility is increased and the beneficial effects of PT are reduced or even reversed.

3.1.2 Optimal policy under commitment

The standard argument made in favour of discretion is that, in practice, "no major central bank makes any kind of binding commitment over the course of its future monetary policy" (Clarida et al. 1999, p. 1671). It is far from obvious, however, that discretion provides a better description of the real-world behavior of central banks than commitment – even to those who have had considerable

[6] The discrepancy between the results in Steinsson (2003) and those in Röisland (2006) and Gaspar et al. (2007) arises because the social loss function differs in the two models. In the latter case the loss function remains as in Equation (3), except that the squared term in inflation is replaced with a squared term in the quasi-difference of inflation (see Roisland, 2006). In the former case analysed by Steinsson (2003), the relative weights in the loss function change and there are additional squared terms in the lagged output gap and the change in inflation rate.

[7] Blake et al. do not address the issue of whether welfare-superior equilibria could be selected under certain delegation schemes in their paper.

input in the policy process (see eg Blinder, 1997). Indeed, commitment may well provide a better description of the behavior of central banks because it rules out systematic policy mistakes which are repeated ad infinitum under discretion. Moreover, as is now familiar in the context of formal 'foward guidance', real-world central banks do commit to contingent rules in which there are binding forward targets for inflation and output; and, as in our theoretical models, these rules commit the central bank except in respect of its contemporary response to shocks. Commitment solutions are also helpful because they provide guidance on the optimal conduct of monetary policy, which adds to our understanding of why some monetary policies perform better than others. Consequently, many papers in the literature have used commitment to shed light on the relative merits of PT from a stabilization perspective. This commitment-based literature has two different strands: the first focuses on the features of optimal policy under commitment; the second considers the performance of simple Taylor-type rules.

Optimal commitment One well-known result from the first strand of literature is that, in the standard New Keynesian model, optimal commitment implies price stationarity and history dependence (Clarida et al., 1999; Woodford, 2003). As emphasised by Vestin (2006), both of these features are consistent with a PT regime.[8] By contrast, an IT regime is inconsistent with price stationarity because it implies base-level drift in the price level in response to inflationary shocks; and it lacks history dependence because 'bygones are bygones.' As discussed by Woodford (2000), history dependence is a robust feature of optimal policy in forward-looking models. However, recent research has shown that the price stationarity result is not robust to minor modifications of the baseline New Keynesian model. For instance, Steinsson (2003) shows that base-level drift in the price level is optimal under commitment if one makes the additional assumption that some fraction of price-setters are backward-looking and follow a simple rule-of-thumb when setting prices. More recently, Levin et al. (2010) show that optimal policy involves considerable base-level drift in response to contractionary demand shocks when the zero lower bound on nominal interest rates is an occasionally-binding constraint.

Even in the absence of the zero lower bound, the price stationarity result is rather fragile. Amano et al. (2012) show that base-level drift is in fact optimal in the standard New Keynesian model if the central bank and agents must make decisions before current shocks to the economy are observed. Intuitively, the benefits of price stationarity in the standard model come from the expectation that policy will offset inflationary shocks, which in turn dampens the impact of shocks on current inflation. But if the central bank cannot observe current shocks to inflation, it cannot react directly to these shocks and so cannot influence price-setters' expectations in a favourable way. The benefits of

[8] We can conceive of cases where history dependence could be destabilizing and move the economy further away from the optimal commitment solution. However, Vestin (2006) shows that the history dependence introduced by PT is beneficial from a welfare perspective. For example, he shows that PT can exactly replicate the optimal commitment policy in the baseline New Keynesian model when there is no persistence in the cost-push shock u_t.

price stationarity are therefore lost and it becomes optimal to treat inflationary shocks as bygones.

A second example is provided by Gerberding et al. (2012), who consider a two-sector extension of the benchmark New Keynesian model with nominal rigidities in both the intermediate goods sector and the final goods sector. They show that price-level stationarity can be very costly in such an environment because optimal commitment implies substantial base-level drift in both sectors, and extending the target horizon under PT lowers but does not eliminate these welfare costs.[9] Intuitively, PT performs poorly in this environment because the price-level target relates to the price level in the final goods sector, which implies that the central bank must accept strong output fluctuations and a suboptimally high degree of base-level drift in the intermediate goods sector to return the price level of final goods to target at the appropriate horizon.

Commitment to Taylor rules The second strand of commitment literature has assessed the performance of IT and PT Taylor-type rules in New Keynesian models. Simple interest rate rules are motivated by the argument that they may provide a better representation of the behavior of real-world central banks, since they are easy to communicate; more robust across alternative models than fully optimal policies (see eg Taylor, 1999); and easier to implement as they do not imply a policy response to economic shocks which cannot be easily observed – see Equation (6) above and Equation (3.6) of Clarida et al. (1999).

Whereas IT Taylor rules relate the nominal interest rate to a measure of the output gap and inflation (like the original Taylor rule), PT Taylor rules respond to a price-level gap rather than the inflation gap:[10]

$$R_t = c_p(p_t - p_t^*) + c_x x_t \tag{8}$$

where $c_p, c_x > 0$ are reaction coefficients, p_t is the log of the aggregate price level and p_t^* is the target price level.

The Taylor rule literature contains several analyses of PT in open economies. An early contribution to this literature was Batini and Yates (2003). They show that the degree of openness of an economy is generally important for comparisons of IT and PT. In their model the real exchange rate enters the Phillips curve, so that the variability of inflation depends on fluctuations in the real exchange rate, with this channel becoming more important as the degree of openness of the economy is increased. On the one hand, PT can have a positive impact on inflation stabilization because the uncovered interest parity (UIP)

[9] Formally, the target horizon is the number of periods in which the price level must be returned to target.

[10] When comparing across regimes using Taylor rules, the reaction coefficients are typically chosen to minimize the social loss function. In addition, some papers in the literature allow for 'interest rate smoothing' through dependence of the nominal interest rate on its lagged value, R_{t-1}.

condition introduces an additional channel through which managing expectations matters for economic outcomes. On the other hand, PT could lead to a deterioration in stabilization relative to IT, because it makes interest rates somewhat more volatile, and this feeds back to greater real exchange rate volatility by the UIP condition. In ranking IT and PT regimes in welfare terms, it is the relative size of these two effects that is important.

Much of the subsequent research in this literature has been carried out by or in conjunction with the Bank of Canada. For instance, Coletti et al. (2008) compared IT and PT in the two-country IMF Global Economy Model (GEM), a medium-scale DSGE model designed to enable open-economy issues to be investigated within a representative-agent framework suitable for policy analysis (see Laxton, 2008). This model contains several sources of nominal rigidity and is forward-looking; it can therefore be viewed as an extension of the baseline New Keynesian model. Coletti et al. calibrated the model for Canada with the US as the second country, giving the BoC-GEM model. They found that a PT Taylor rule outperforms an IT one in terms of inflation and output gap volatility, primarily because shocks to the terms of trade strengthen the case for PT due to its role as a nominal anchor that stabilises the domestic price level.

Dib et al. (2008) investigated the impact of PT within a medium-scale open economy New Keynesian model whose parameters are estimated on Canadian data. The model is augmented with credit frictions as in Bernanke et al. (1999), and entrepreneurs enter into one-period nominal debt contracts in order to finance investment. In this model, PT lowers the distortion in the economy due to nominal debt contracts, because inflation expectations are better stabilized than under IT. Real risk faced by entrepreneurs is therefore reduced, so that resources are allocated more efficiently, and this increase in efficiency means that nominal interest rates need not need to vary as much in order to minimize the standard welfare loss associated with nominal price rigidity, so that the real interest rate volatility is also reduced. PT also tends to outperform IT in small open economies with multiple sectors which are subject to sector-specific terms of trade shocks, because the gains from PT due to extra stabilization through the expectations channel outweigh the losses involved in responding to additional shocks so as to stabilize the aggreagate price level (see Murchison, 2010 for a brief review).

In summary, when monetary policy commits to Taylor-type rules, PT tends to dominate IT in several different variants of open economy New Keynesian models. Hence, the main results highlighted by Giannoni (2010) in the context of the baseline (closed-economy) New Keynesian model – namely that PT Taylor rules outperform IT ones in welfare terms and are more robust – largely appear to carry over to richer models.[11] By contrast, optimal commitment results in favour of PT appear to be rather fragile to minor modfications of the baseline

[11] It is important to note however that dominance of PT over IT is not a necessary outcome when policymakers commit to Taylor-type rules. For example, a recent working paper by Coletti et al. (2012) which considers an extended version of the BoC-GEM model with separate energy and non-energy sectors finds that the impact of PT is ambiguous. PT improves

model. We argue, however, that it is the former case that is of most relevance for real-world central banks, because fully-optimal policies do not perform well across a range of alternative models of the economy and are more difficult to implement in practice. We return to the argument that PT Taylor rules will tend to dominate IT Taylor rules after discussing whether this result also occurs in the presence of the zero lower bound on nominal interest rates and financial market frictions.

3.2 The zero lower bound on nominal interest rates (ZLB)

In simple terms, the ZLB states that nominal interest rates cannot fall below zero in an economy where money is untaxed and can be stored without cost. If it did, so the argument goes, rational agents would have no incentive to hold nominal government debt, because it would be dominated by money, a liquid asset with a fixed nominal return of zero. Aided by recent advances in computing power, the macro stabilization literature has incorporated the ZLB into New Keynesian models as an occasionally-binding constraint, that is, a constraint of the form:[12]

$$R_t \geq 0, \qquad \forall t \tag{9}$$

In these analyses, economic agents take the ZLB into account when forming expectations about the future course of monetary policy, and the ZLB binds only a small fraction of the time – usually between zero and 10 per cent of all simulated periods. Eggertsson and Woodford (2003) was the seminal contribution in this literature. They introduce the ZLB constraint into the baseline New Keynesian model and solve for the optimal commitment policy. In addition, they also evaluate the performance of various monetary policies at the ZLB using numerical analysis. Eggertsson and Woodford's most important findings were as follows.

First, the optimal commitment policy with a ZLB takes the form of a state-contingent PT targeting rule where the target price level is updated every period based upon the previous period target shortfall. Intuitively, the optimal policy includes a state-contingent target price level so that inflation expectations receive extra stimulus at times when deflationary pressure is greatest; in turn, this implies that real interest rates are lowered sharply as a ZLB episode worsens, boosting output. Since this policy implies a state-contingent price level at times when the ZLB is binding (ie at times when the target shortfall is non-zero), it in fact implies base-level drift in response to shocks. Consequently, the optimal commitment policy with the ZLB contains features of both IT and PT, in contrast to the baseline New Keynesian model.

Second, Eggertsson and Woodford show that the state-contingent optimal policy – which would probably be difficult to implement in practice – can be

stabilization in the case of non-energy commodity supply shocks, but IT does slightly better in response to energy commodity supply shocks, as well as producing somewhat better outcomes in response to commodity demand shocks.

[12] For a review of early studies on the zero lower bound, see Yates (2004).

approximated well by a simple interest rate rule that aims at a fixed target price level (implying zero base-level drift), while a standard IT-type interest rate rule performs very poorly by comparison. The key to this result is that under IT, the real interest rate does not respond to the severity of a ZLB episode, because inflation expectations are fixed on a constant inflation target. Under the PT rule, by contrast, the real interest rate falls as the severity of ZLB episodes worsens, because agents rationally expect future inflation to return the price level to target. This stimulus to the output gap from lower real interest rates aids the economy's recovery from the ZLB.

Formal welfare analyses of optimal policy at the ZLB have been conducted by Adam and Billi (2006, 2007) and Nakov (2008). The model in Adam and Billi (2006) follows Eggertsson and Woodford (2003) but they calibrate the model to US data using shocks from the Great Moderation period. In order to evaluate social welfare, these shocks are used to simulate the model, which is solved using numerical methods. They find that zero nominal interest rates occur rather infrequently under the optimal commitment policy – only about one quarter in every 17 years, or an unconditional probability of 1.5 per cent. As a result, the additional welfare loss due to the ZLB is quite small at approximately one per cent of the welfare loss generated by sticky prices alone, though this estimate is quite sensitive on the upside to the variance of innovations to the natural rate of interest. Adam and Billi (2007) extend the analysis of optimal policy in the baseline New Keynesian model to the case of discretion. They find that the welfare losses imposed by discretion relative to optimal commitment increase by around two-thirds when the ZLB is an occassionally-binding constraint, suggesting that the potential welfare gains from well-designed alternatives to IT are likely be somewhat larger when one accounts for the ZLB in New Keynesian models.

Nakov (2008) takes up this theme. His contribution is to assess social welfare for a variety of *zero-truncated* Taylor-type interest rate rules, which he argues provide a more plausible representation of real-world monetary policy than optimal policies. Using the standard New Keynesian model, he assesses the performance of IT and PT Taylor rules. All the rules perform poorly relative to optimal commitment: the social losses for the PT and IT Taylor rules are 800 and 1400 per cent respectively, with the loss under IT being much larger if interest rate smoothing is not permitted. On some level this finding is not surprising: Taylor rules rarely perform well in individual models of the economy but have the advantage of robustness across models (Taylor, 1999). Crucially, however, constraining the IT-PT comparison to simple Taylor-type rules leaves intact the conclusion that PT is beneficial in the presence of the ZLB.

Finally, an important recent contribution to the ZLB literature is Coibion et al. (2012). They focus on a New Keynesian model with an occasionally-binding ZLB and simple Taylor-type rules. Crucially, they allow for a positive steady-state rate of inflation around which the model is log-linearized, so that they can study the optimal rate of inflation in the presence of the ZLB. Under

IT, moderate trend inflation is optimal in order to reduce the probability of hitting the ZLB and the welfare losses associated with this. The welfare gains of PT are magnified relative to the case where steady-state inflation is not optimized, because the likelihood of the ZLB being reached is much lower for any given steady-state inflation rate under a PT regime than an IT one. As a result, the welfare benefits of lowering steady-state inflation can be realized while maintaining a low frequency of ZLB episodes. The potential welfare gains from PT through this channel are quantitatively quite significant: raising the response to the price level in the Taylor rule from zero to 1/4 implies a welfare gain equivalent to a permanent increase in aggregate consumption of 0.5 per cent.[13] This finding, and the results of Nakov (2008), are further examples of PT dominating IT when policymakers commit to simple Taylor-type rules.

3.3 Financial frictions

An issue brought to the fore by the recent financial crisis is the potential importance of financial market imperfections for the operation of the macroeconomy. Covas and Zhang (2010) focus on the implications of financial market imperfections for comparisons of IT and PT. To do so, they compare IT and PT in a New Keynesian model augmented with financial market imperfections in both debt and equity markets. The structural parameters of the model are estimated for the Canadian economy. They find that PT outperforms IT in terms of stabilization because the expectations channel means that inflation is better anchored under PT, so that it is less costly for the central bank to address financial market distortions through monetary policy. It should be noted, though, that the benefits of PT are smaller in the presence of financial market imperfections and fall as the importance of financial market frictions is strengthened.

Bailliu et al. (2012) study the interaction between macroprudential rules and monetary policy in a model with financial market imperfections. As in Covas and Zhang (2010), financial frictions are introduced through debt contracts à la Bernanke et al. (1999), implying an inverse relationship between the external finance premium and firm net worth. The model is estimated on Canadian data and calibrated accordingly. Under optimised IT and PT Taylor rules, it is beneficial to respond to financial imbalances in both a Taylor rule and a macroprudential rule. Nevertheless, PT rules still delivers substantial relative welfare gains relative to IT because it is history dependent. These results provide additional confirmation that augmenting New Keynesian models with financial frictions does not substantially alter the relative performance of IT and PT regimes, as well as additional evidence that PT tends to dominate IT when policymakers commit to simple to Taylor-type rules.

[13] Positive trend inflation has three distinct costs in New Keynesian models; see Ambler (2007) for a discussion. The traditional welfare cost of inflation due to inflation acting as a tax of money holdings (see Bailey, 1956) is not one of them. It is therefore conceivable that the welfare gains attainable from lowering trend inflation under PT could be larger than estimated by Coibion et al. (2012).

Lastly, Dib et al. (2013) also find that PT Taylor rules outperform IT ones. They consider the same model as in Dib et al. (2008) – ie a medium-scale open economy New Keynesian model augmented with credit frictions as in Bernanke et al. (1999), entrepreneurs who enter into one-period nominal debt contracts to finance investment, and financial shocks. The parameters of the model are estimated on Canadian data. The main finding is that a PT Taylor rule outperforms an IT Taylor rule in welfare terms when the Taylor rule is extended to include a response to the external finance premium and regardless of whether interest rate smoothing is permitted or not. The key to this result is that the welfare costs of the nominal debt distortion are linked to real interest rate variability, which is reduced under PT because large movements in nominal interest rates are not necessary to stabilize the economy due to history dependence. Consequently, introducing nominal debt contracts and financial shocks strengthens the case for PT.

3.4 Summary

A key finding from the macro stabilization literature is that PT has beneficial effects in New Keynesian models. These microfounded models have become the workhorse for monetary policy analysis in recent years. In these models, optimal policy implies price stationarity and history dependence – key features of a PT regime. Although the optimality of price stationarity is not robust to minor modifications of the baseline New Keynesian model and history dependence need not guarantee that PT dominates IT, when monetary policymakers commit to simple Taylor-type interest rate rules, PT consistently outperforms IT in welfare terms. This result is robust in the sense that it holds for several extensions of the baseline New Keynesian model, including open economies, the zero lower bound, and financial market imperfections.[14] The beneficial effects of PT therefore appear to hang on the joint hypothesis that agents are rational and the economy New Keynesian. Accordingly, we focus below on empirical evidence on rational expectations and the performance of New Keynesian models.

4 Survey and experimental evidence on rational expectations

As our discussion above makes clear, the stabilizing properties of expectations under PT hang on the assumption that the regime is fully understood by forward-looking agents with up-to-date information – ie the rational expectations assumption. It is therefore crucial to test this scientifically. Broadly speaking, evidence on rational expectations has come from two lines of enquiry: studies investigating the empirical performance of aggregate macro models with

[14] This appears to be the case because – unlike price stationarity – history dependence is a robust feature of optimal policy in New Keynesian models. See Giannoni (2010) for a discussion of history dependence in the context and IT and PT Taylor rules.

rational expectations; and surveys and experiments. We consider the second line of enquiry in this section.

4.1 Survey evidence on rational expectations

Surveys are a popular way of testing the rational expectations assumption. At an intuitive level this popularity is hardly surprising: surveys give researchers the freedom to 'ask the right question' while also enabling them to collect expectations data at the micro level at which household decisions are taken. Tests of rational expectations using individual-level survey data can also avoid biases that arise from aggregation or pooling of forecast data, as in the case of consensus forecasts (see Bonham and Cohen, 2001).

Some well-known surveys are the Livingston Series in the US and the Confederation of British Industry survey in the UK (see Carlson and Parkin, 1975). Early studies using data from these surveys concluded that expectations are not strictly rational, because they do not satisfy the efficiency property: information available to survey participants was not independent of expectation errors, though some studies in the literature were not able to reject efficiency conclusively (see Holden et al. 1985, Ch. 3). However, there are some well-known difficulties with the survey approach. In markets the expectations that matter are those of the active market participants and surveys may not identify these people: they question people who may not be active. This in turn means that survey respondents may be inattentive and poorly informed. Other potential problems include truthfulness and accurate recall.

In addition, an important recent contribution to the survey literature is made by Andolfatto et al. (2008). They show that when the true data-generating process (DGP) is a DSGE model with switches in monetary policy regime and agents who form expectations rationally using the Kalman filter, standard regressions of the kind used in the survey literature on inflation expectations will incorrectly reject the rational expectations hypothesis. This finding suggests that several recent papers reporting biased inflation expectations – eg Thomas et al. (1999) and Mankiw et al. (2003) – cannot be interpreted as conclusive statistical evidence against rational expectations.[15] In short, the problem is that these papers are subject to small-sample bias which is exacerbated if short run learning dynamics are generated by shifts in monetary policy regime. Given this and the other potential problems with surveys outlined above, it is important to bring other sources of evidence to bear on whether expectations are rational.

4.2 Experimental evidence on rational expectations

An alternative approach to surveys that has gained popularity in recent years is experimental economics. Like surveys, experiments are flexible and can be targeted at individuals. The main advantage of the experimental approach is

[15] In particular, these papers report evidence that inflation expectations in the data are not mean unbiased, implying that forecast errors are serially correlated.

that the behaviour of participants can be studied within a particular economic context chosen by the researcher. Consequently, this approach has been widely used in the microeconometric literature to evaluate the effectiveness of proposed policy interventions, though more sparingly in macroeconomics (see Angrist and Pischke, 2010). An early contribution to the literature was Smith et al. (1988). They study spot asset trading in an experimental environment and conclude that traders' expectations converge upon rational expectations as they acquire additional experience; in their model this eventually leads to speculative bubbles being ruled out. However, the findings of other studies in the literature are generally more mixed. For instance, Bloomfield and Hales (2002) found that experimental subjects who were told that the data-generating process was a random walk did not have static expectations, the rational expectations solution in this case. Their beliefs were instead consistent with switches between a trending regime and a mean-reverting one.

In relation to the IT-PT debate, Amano et al. (2011) investigated whether experiment participants were able to accurately forecast inflation in a simulated PT monetary policy regime. They found that participants' forecasts did adjust under PT, but not to the extent that the implications of targeting the price level were fully reflected in their expectations. It is important to note, however, that these findings were obtained under conditions relatively unfavourable to PT: participants were given minimal information about the model economy and had the shift in monetary policy regime explained to them only once. It seems likely that agents' expectations would take account of the implications of PT more fully if they had more information about the economy or ongoing communication explaining the new regime, though it is not clear whether such a strategy would close the gap completely. It is also worth pointing out that inflation was forecast at the individual level in the study, yet in practice many economic decisions are taken at the level of the household. Since the more well-informed agents in the household will presumably take responsibility for economic decisions, focusing on understanding of PT at the individual level may understate the true level of public understanding. Indeed, as we argue below, market outcomes are likely to be dominated by the actions of well-informed agents, regardless of whether we are considering individuals or household units.

One weakness of the experimental literature is that it is subject to some of the same criticisms as surveys: subjects may deliberately alter their behaviour under experimental conditions; and experiments cannot identify active market participants unless they are aimed at a specific group, which in turn may reduce the extent to which they can shed light on macroeconomic issues. It is also well-known that experimental outcomes are highly sensitive to the 'rules of the game', a point highlighted by John List in a series of papers. Lastly, the usefulness of experiments in addressing macro issues is likely to be limited by the cost of large-scale experiments and the fact that it is important but difficult to provide experimental participants with the same flow of information that they would have access to outside of an experimental environment. Given the potential problems with surveys and experiments, we believe that it is preferable to test macro theories at the aggregate level using macro data when such results can be established on strong statistical grounds.

5 Testing the models used in evaluating price-level targeting

When as here we are concerned with the macro implications of a new policy regime, it is preferable to test our macro models of policy on macro data. This approach in itself is not new but we proceed, after a brief review of the current literature, to a discussion of more recent developments that have moved the literature in the direction of statistically testing the restrictions implied by New Keynesian models where consumers and firms have rational expectations.

5.1 The empirical performance of macroeconomic models

Following the rational expectations revolution of the 1970s, macroeconomic models with rational expectations became commonplace, yet few formal attempts were made to test these expectations against alternative expectational assumptions. Fair (1993) is an early exception. He sets up a model of the economy which nests both rational and adaptive expectations, making a direct test of the rational expectations hypothesis possible. The estimated equations come from a medium-scale macro model augmented with current and lagged values of macro variables which are considered potential determinants of the dependent variables, as well as future expected values based on current information. The test of rational expectations then amounts to a test of the joint significance of the coefficients on the forward-looking variables. In total, this test is conducted for 16 separate equations, including 7 describing the behaviour of the household sector, 5 describing the behaviour of firms, and additional equations for investment, employment, interest rates, and asset prices. Each equation is estimated using quarterly data over the period 1954 to 1986.

The results show some support for the rational expectations hypothesis. In particular, 8 of the 16 equations estimated have significant lead coefficients at the 1% level for at least one of four alternative tests for each equation, with the strongest support coming from the 7 household equations, of which 5 are statistically significant. In the other 8 cases, however, the null hypothesis of adaptive expectations is not rejected. This testing procedure is not without problems, though, as recent research has highlighted. One important problem is that the model solution for the forward expectations of inflation will importantly include lagged inflation; hence it is unclear that the tests have much power or indeed that the relevant equations are identified. To give a pertient example, it is widely recognized that equations with lagged inflation may be equivalent to the solution of those with future expectations.[16] A second problem is that the

[16] For example, the solution for inflation under commitment in the baseline New Keynesian model is $\pi_t = \delta\pi_{t-1} + \phi\Delta u_t$ (Clarida et al. 1999, p. 1704). Hence $E_t\pi_{t+1} = \delta\pi_t + \phi E_t\Delta u_{t+1}$

test is a partial information test: each equation is tested on a piecewise basis, so there is no unambiguous accept/reject decision for the model as a whole. More importantly, this means that the cross-equation restrictions implied by rational expectations are left untested. More recent literature that we discuss below has overcome these weaknesses by using full-information methods. A final weakness of the test conducted by Fair (1993) is that it is vulnerable to the Lucas critique of policy evaluation (Lucas, 1976). In particular, the model is not derived from first-order conditions for households and firms and so does not contain 'deep parameters' that can be expected to be invariant to structural change; nor is the set of equations chosen for testing disciplined by rigorous microfoundations.

In order to address these concerns, macroeconomic models are now built, as standard, from models of the economy with market clearing and optimizing households and firms with rational expectations. These micro-founded dynamic stochastic general equilibrium (DSGE) models were popularized by the pioneering contribution of Kydland and Prescott (1982), who showed that, in such a model, technological shocks alone could account for a surprisingly large fraction of US output volatility. Indeed, as Fernández-Villaverde (2010, p. 4) observes, "the amazing feature was how well the model did despite having so little of what was traditionally thought of as the necessary ingredients of business cycle theories: money, nominal rigidities, or non-market clearing." Since Kydland and Prescott's 1982 paper, researchers have augmented real business cycle models with New Keynesian nominal rigidities in order to provide a plausible stabilization role for monetary policy. These models do a good job of matching several key features of aggregate data when hit with both real and nominal shocks, as we explain below. It is this combination of theoretical foundations and empirical performance that has made New Keynesian models popular in the macroeconomic stabilization literature discussed in Section 3.

An important contribution to the applied New Keynesian literature was Christiano et al. (2005). They set up a medium-scale New Keynesian model with staggered wage contracts, variable capacity utilization, investment adjustment costs, and indexation of prices and wages to past inflation. The model parameters were estimated on quarterly U.S. data over the postwar period using Bayesian methods. The estimated model was able to replicate the dynamic patterns of the impulse response functions of several important macroeconomic aggregates in response to monetary policy shocks (i.e. an unexpected fall in the Federal Funds rate), as estimated using a vector autoregression (VAR). In particular, the model does a good job of accounting quantitatively for the impulse responses of inflation, output and the real wage, as well as the lagged, hump-shaped responses of consumption, investment, profits and labour productivity in the data. Therefore, augmenting the New Keynesian model with real and nominal frictions enables it to mimic the dynamic responses of several key variables to nominal shocks. Smets and Wouters (2007) go even further. They show that a Bayesian-estimated New Keynesian model with several sources of

is a function of π_{t-1}. Similar arguments apply to the output gap solution.

real and nominal rigidities and 7 orthogonal shocks can match several additional dynamic features of U.S. data, and that the model as a whole performs well in out-of-sample forecast tests. More specifically, the estimated model does as well as a BVAR(4) in terms of out-of-sample forecast performance using the marginal likelihood criterion; and a more traditional out-of-sample forecasting exercise shows that the model performs comparably to a BVAR(4) over short horizons up to one year, and considerably better at longer horizons such as 2 or 3 years.

Some recent papers have considered refinements of rational expectations in medium-scale New Keynesian models in the spirit of Christiano et al. (2005) and Smets and Wouters (2007). Schmitt-Grohé and Uribe (2012) show how one can model 'news' in a DSGE framework, that is, the arrival of information on future shocks prior to the time when those shocks are actually realized. Fujiwara et al. (2011) and Khan and Tsoukalas (2011) estimate medium-scale DSGE models with news shocks and sticky prices and wages. Fujiwara et al. (2011) find that the inclusion of news shocks significantly improves the fit of the New Keynesian model in the case of the US economy but not Japan; and Khan and Tsoukalas (2011) show that the importance of news shocks is sensitive to model structure, shocks, and the data used in estimation – making it hard to draw any general conclusions about the contribution of news shocks to business cycles. Alternative refinements to rational expectations that have had some success empirically include narrowing the information set of agents (ie imperfect information) as in Levine et al. (2012); 'sticky information' as in Mankiw and Reis (2002); and constraints on information flow due to 'rational inattention' as in Mackowiak and Wiederholt (2011).[17] The main attraction of these approaches is that they can help New Keynesian models to match inertia in the data without the need for features such as price indexation and habit formation whose empirical basis is questionable. These other approaches have not yet been shown to convincingly outperform standard models with fully rational expectations, but they are a promising avenue for future research.[18]

In summary, New Keynesian models seem to do a reasonably good job of matching dynamic features of aggregate macro data as summarized by VARs,

[17] In addition, some recent papers have considered more drastic deviations from rational expectations such as behavioral expectations (as we discuss below) and adaptive learning (Evans and Honkapohja, 2001; Slobodyan and Wouters, 2012). See Milani (forthcoming) for a survey of approaches to modeling expectations in DSGE models.

[18] Levine et al. (2012) estimate an NK model that nests rational expectations and behavioral expectations using Bayesian methods. They find that the 'composite' (partly behavioral) expectations model outperforms the model with rational expectations, in contrast to the results we present in Section 5.4 based on the method of indirect inference. However, as they note (p. 1298), "a limitation of the likelihood race methodology is that the assessment of model fit is only relative to its other rivals with different restrictions. The outperforming model in the space of competing models may still be poor (potentially misspecified) in capturing the important dynamics in the data." It is simply not clear what one is to conclude from a relative ranking of models, if all of them are rejected by the data behaviour (as may be the case here). An advantage of indirect inference is that it provides both a relative ranking of different models and an absolute test of whether any particular model is rejected by the data.

and the literature has recently considered refinements to rational expectations that could improve the fit of these models in future. The next logical step is to test models with rational expectations against the data using formal statistical tests that accept or reject the baseline model and variants of it. Such tests should narrow down the set of models that need to be considered in future research by rejecting particular models (or variants) that clearly cannot match the data. As we discuss below, this challenge has been taken on by a recent strand of applied macro literature that exploits the use of VARs as a description of the time-series properties of aggregate macro data. In light of the promising results dicussed above, this recent literature has focused on New Keynesian DSGE models.

5.2 Testing DSGE models: indirect inference

Indirect inference is a simulation-based method used for estimating or evaluating economic models. Its distinguishing feature is the use of an 'auxiliary model' – which need not be correctly specified – to represent the time series properties in the data. Estimation and evaluation are then based upon the auxiliary model, which acts as a criterion function that selects important features in the data; this is the sense in which inference is 'indirect'. Indirect inference has most commonly been used to estimate structural economic models (Smith, 1993; Gourieoux and Monfort, 1996). An important advantage of this approach is that estimation is possible when the likelihood function is difficult to evaluate or analytically cumbersome (Canova, 2007). Indeed, indirect inference can be used to estimate almost any economic model from which data can be simulated. Indirect inference basically chooses the parameters of the macroeconomic model so that, from the point of the view of the auxiliary model, the actual and simulated data look similar. In this respect, researchers can use indirect inference to focus upon matching those aspects in the data which they view as most important. The extension of indirect inference to the case of evaluating macroeconomic models is discussed in detail below, as this our main focus in this section.

It is worth mentioning at this point that there have been well-known problems with evaluating DSGE models of the economy. Hansen and Heckman (1996) raised two problems which are relevant for our discussion here. The first is that calibrating macroeconomic models is not a well-defined scientific procedure: researchers can end up calibrating the same models with very different parameters based upon their interpretation of the empirical literature. The second issue relates to the way that the empirical performance of these models is evaluated – typically, a small (but not precisely defined) set of simulated moments was to be compared informally with the same moments in the data to judge whether the model was a success. As should be clear from the discussion of the previous section, the most influential papers in the recent literature are those which have responded to the first challenge hands-on by estimating macro models on the data using full-information methods. The second issue raised by Hansen and Heckman is very much alive today, however. In fact, as noted above, the time series properties of estimated DSGE models are commonly compared to alternative representations such as reduced form VARs (or their implied impluse response functions) in much the same way that the business cycle moments of early DSGE models were subjectively compared with the business cycle moments in the data. Something stronger is needed to choose between alternative models of the economy – a formal statistical test that will accept or reject the theory. The method of indirect inference provides such a test.

Indirect inference can be used to test whether a DSGE model can simulate behaviour that is 'like' the behaviour in the data, where the data behaviour is summarised by the reduced form representation of some unknown true model, or an approximation to it. Typically, the reduced form approximation is a VAR or a VECM.[19] Then the question is whether the structural DSGE macro model could have been the generating mechanism of these coefficients. We can answer this question by simulating the DSGE model over the same data period with repeated samples of its own errors – these samples will give us 'pseudo-histories' that the model was capable of creating, and we can then ask whether the actual history (as captured by its estimated reduced form coefficients) could have been one of these histories.

To formalize this intuition, it is worth briefly considering the indirect inference testing procedure. The method works as follows. Suppose we have a DSGE model with a fixed vector of structural parameters θ which can be taken as given, having been reached either by estimation or calibration. Let the vector of auxiliary VAR parameters associated with simulated data of length T periods from the DSGE model be denoted $\alpha_T(\theta)$. The corresponding parameter vector from the VAR on the actual data of lengh T periods is denoted a_T. Under the null hypothesis that the DSGE model is correct, the Wald statistic $W(\theta)$ for a test of the model against the data is based on the difference between the VAR parameters estimated from the data a_T, and the mean VAR parameter vector $\overline{\alpha_{T,N}(\theta)}$ estimated from N bootstrapped samples from the DSGE model:[20]

$$W(\theta) = d'\Omega(\theta)^{-1}d \tag{10}$$

where $d \equiv a_T - \overline{\alpha_{T,N}(\theta)}$ and $\Omega(\theta)^{-1}$ is the inverse variance-covariance matrix of the distribution of d.

Notice that the Wald test statistic can simply be interpreted as a quadratic loss function in the deviations of the VAR parameters estimated on the data

[19] Wickens (2011, pp. 506-8) has shown that DSGE models will have a VARMA or VARIMA representation. Indirect inference can proceed with approximations to this because the researcher runs a VAR (or VECM) using the data from the simulated model; hence the model's simulations of the approximation give rise to the distribution against which the approximating VAR or VECM on the data is tested for whether it comes from this distribution.

[20] Bootstrapping N samples avoids the weaknesses associated with a test based on a single random sample of simulated data.

from those implied (on average) by the bootstrapped DSGE model; as such it is unambiguously greater than zero. The weights attached to each parameter will generally depend on the joint uncertainty about all estimated parameters in the VAR,[21] but those parameters whose estimates vary less from sample to sample will be given a higher weight in the test. Essentially the Wald statistic is assessing whether the VAR parameters from the data fall outside the joint confidence limit of the model-simulated parameter distribution. High values of the Wald statistic will reject the null hypothesis that the DSGE model is the data-generating process (DGP) because they tell us that it is unlikely the VAR parameters in the data could have been generated by the simulated model.

Finally, it should be noted that the Wald statistic has typically been compared to its simulated bootstrap distribution and not the asymptotic distribution implied by theory under the null. It is therefore the percentiles of the bootstrap distribution that provide the basis for accepting or rejecting a particular model, with a Wald statistic higher than the 95% critical value being the standard rejection criterion. The Wald test is thus an example of what Canova (2007) calls a 'size test' of an economic model. In a simple univariate size test the researcher might compare (say) a particular correlation in the data with the distribution of correlations implied by many simulations of an economic model, rejecting the model if the real-world correlation does not fall within a 95% interval.[22] The Wald test is based on the same underlying logic, but it is more general because it enables the researcher to conduct an overall test of the model based on its joint parameter distribution. Moreover, it is important to note that although the Wald statistic in Equation (10) considers only a direct test of the VAR parameters, the test can easily be augmented to include (say) the variances of key macro variables or the impulse response functions implied by the VAR coefficients. The papers we discuss below take this more general approach, so we provide some technical details in a footnote for interested readers.[23]

5.3 Testing the degree of nominal rigidity in the US economy using indirect inference

Le et al. (2011) use indirect inference to investigate the degree of nominal rigidity in the US economy in the postwar period. This paper provides a useful example of indirect inference and is relevant to our discussion here, because it sheds light on the performance of New Keynesian models in formal statistical tests. Le et al. consider a weighted version of the Smets and Wouters (2007) US model in which a New Classical sector with flexible wages and prices and

[21] The reasoning is simply that a DSGE model implies a joint parameter distribution, because it is a set of simultaneous equations.

[22] For example, Lim and McNelis (2008, Ch. 9, pp. 171–72) simulate the distribution of the real-wage employment correlation in an open economy New Keynesian model and discuss how a size test could be conducted using this distribution.

[23] If we let $g(.)$ denote a vector valued function that includes the different aspects of the data that are to be tested, the Wald statistic will be amended to $W(\theta) = d'\Omega(\theta)^{-1}d$, where now $d \equiv g(a_T) - \overline{g(\alpha_{T,N}(\theta))}$.

a one-quarter information lag are introduced. The coefficients of the model – including the relative weight on the New Classical sector – are chosen to minimize the Wald statistic in the VAR parameters and the variances of key macro variables. Several notable results emerge. The most striking results come from the post-1984 Great Moderation period: the Wald-minimizing model has a weight of almost 1 on the New Keynesian wage and price sectors and passes the Wald test comfortably with a percentile of 83.8%.[24] These results confirm that New Keynesian models with rational expectations can successfully mimic some key features of the data *and* pass a stringent statistical test against macro data. Interestingly, when the model is estimated on the entire postwar period the Wald-minimizing weights on the New Keynesian part of the model are noticeably lower (suggesting that nominal rigidity has increased over time), and the minimum Wald of 98.7% exceeds the standard 95% cut-off value, so that the model is rejected, although it would pass a more lenient test at 99%. Overall, the results of Le et al. provide support for New Keynesian models while suggesting, at the same time, that there is scope to improve upon the benchmark model. In the next section we ask whether a behavioral expectations version of the New Keynesian model can do so.

We focus on behavioral expectations as an alternative because these models have had some success in accounting for puzzles in the finance literature. For instance, Bernatzi and Thaler (1995) show how 'loss aversion' could account for the equity premium puzzle, while De Grauwe and Grimaldi (2006) show that a behavioral finance model of the foreign exchange market can produce excess kurtosis and fat-tails in exchange rate returns – features which have been documented in numerous empirical studies of financial markets. The behavioral finance approach has also had some success in formal statistical tests against the data. As shown by ap Gwilym (2010), for example, a simple behavioral finance model in the spirit of De Grauwe and Grimaldi (2006) cannot be rejected as the data-generating process (DGP) of the FTSE, based on all its time series properties. One problem, however, is that empirical tests of individual series come up against low power to reject the null that the model is the DGP, as demonstrated by the fact that Meenagh et al. (2007) showed in earlier paper that an efficient markets model incorporating rational expectations and regime-switching could also match the time series properties of the FTSE. This difficulty points to need for higher-power multivariate tests to distinguish between rational and behavioural expectations as the DGP. The macro test we discuss in the next section is a multivariate test that overcomes the problem of low power.

5.4 A test of rational versus behavioral expectations in a macro model of the US

Liu and Minford (2012) test a rational expectations version of the baseline New Keynesian model against a version based on behavioral expectations. The behavioral model is a stylized DSGE model similar to the model in De Grauwe

[24] The standard percentile that must be exceeded for rejection of the model is 95%.

(2010). It includes a standard aggregate demand equation, an aggregate supply function, and a Taylor rule:

$$\widetilde{x}_t = \tilde{E}_t \widetilde{x}_{t+1} - a_1(R_t - \tilde{E}_t \pi_{t+1}) + \varepsilon_{1t} \tag{11}$$

$$\pi_t = b_1 \widetilde{x}_t + \beta \tilde{E}_t \pi_{t+1} + k\varepsilon_{2t} \tag{12}$$

$$R_t = (1 - c_1)(c_2 \pi_t + c_3 \widetilde{x}_t) + c_1 R_{t-1} + \varepsilon_{3t} \tag{13}$$

where $\widetilde{x}_t$ is the output gap, π_t is the rate of inflation, R_t is the nominal interest rate, and ε_{1t}, ε_{2t}, and ε_{3t} are the demand error, supply error and policy error respectively. These errors are assumed to be autoregressive processes. The errors are extracted from the model and the data; thus the model implies the errors, conditional on the data. Equation (11) is the aggregate demand equation, where $\tilde{E}$ refers to expectations that are not formed rationally. The aggregate demand function is standard, including the expectation of output gap in the next period and real interest rate. Equation (12) is a New Keynesian Phillips curve augmented with behavioral expectations. Equation (13) is a Taylor rule with interest rate smoothing.

The difference between the behavioral and rational expectations model lies in expectations formation. The expectation term in the behavioural model, $\tilde{E}$ is the weighted average of two kinds of forecasting rule. One is the fundamental forecasting rule, by which agents forecast the output gap or inflation at their steady state values. The other one is the extrapolative rule, by which individuals extrapolate the most recent value into the future. Thus:

$$\tilde{E}_t^f \widetilde{x}_{t+1} = 0 \tag{14}$$

$$\tilde{E}_t^e \widetilde{x}_{t+1} = \widetilde{x}_{t-1} \tag{15}$$

$$\tilde{E}_t^{tar} \pi_{t+1} = 0 \tag{16}$$

$$\tilde{E}_t^{ext} \pi_{t+1} = \pi_{t-1} \tag{17}$$

Equation (14) and (15) are the forecasting rules for the output gap, while Equation (16) and (17) are the equivalents for inflation. The steady state output gap is zero, while the inflation target in the Taylor Rule is the steady state inflation rate, set at zero since the data is linearly detrended and demeaned.

In De Grauwe (2010), it is assumed that the market forecast is the weighted average of the fundamentalist and extrapolative rules. Equation (18) is the market forecast for the output gap, while Equation (19) is for inflation.

$$\tilde{E}_t \widetilde{x}_{t+1} = \alpha_{f,t} * 0 + \alpha_{e,t} \widetilde{x}_{t-1} = \alpha_{e,t} \widetilde{x}_{t-1} \tag{18}$$

$$\tilde{E}_t\pi_{t+1} = \beta_{tar,t} * 0 + \beta_{ext,t}\pi_{t-1} = \beta_{ext,t}\pi_{t-1} \quad (19)$$

where $\alpha_{f,t}$ and $\alpha_{e,t}$ are the probabilities that agents will use a fundamentalist and extrapolative rule for forecasting the output gap, $\beta_{tar,t}$ and $\beta_{ext,t}$ are the equivalents for inflation. These probabilities sum to one and are determined by past success of the two rules in an intuitive way.

The solution method to the behavioural model is obtained by substituting the expectation formation of Equation (18) and (19) into Equation (11) and (12). The model therefore becomes

$$\widetilde{x}_t = \alpha_{e,t}\widetilde{x}_{t-1} - a_1(R_t - \beta_{ext,t}\pi_{t-1}) + \varepsilon_{1t} \quad (20)$$

$$\pi_t = b_1\widetilde{x}_t + \beta(\beta_{ext,t}\pi_{t-1}) + k\varepsilon_{2t} \quad (21)$$

$$R_t = (1 - c_1)(c_2\pi_t + c_3\widetilde{x}_t) + c_1R_{t-1} + \varepsilon_{3t} \quad (22)$$

As Equations (20)-(22) make clear, the behavioural expectations model is purely backward-looking.

The stylized DSGE model with rational expectation is defined as Equation (11)-(13) except that the expectations are formed rationally. The only specification difference between the two models is in the nature of these expectations. Thus the comparison precisely tests the different specification of expectations, allowing each model the benefit of reestimation of parameter values. The rational expectation version of the model can be solved in the standard way; Dynare (Juillard, 2001) is used for this.

For this model the results, after allowing each model to be reestimated by indirect estimation (in practice this involves a search for the value of the coefficient vector θ that minimizes the Wald), are shown in Table below. The Table shows the Wald percentiles for both models, broken down by variances ('volatility') alone, VAR coefficients ('dynamics') alone, and the full vector of descriptors. In all aspects the behavioural model is strongly rejected, whereas the rational expectations model is easily accepted with high p-values (100 minus the Wald percentile). It must be emphasised that this is after allowing each model to explore all possible values for all the model's parameters, to find the set of parameter values that gets closest to the data behaviour. Each model is then tested on its own 'best' parameters.

It may seem counter-intuitive that a theory of expectations so apparently 'unrealistic' as rational expectations can replicate macro behaviour so much better than apparently 'realistic' alternatives such as behavioural expectations. However, as Muth (1961) notes, in an economy with informed and misinformed agents, the well-informed can profit by selling their superior information to the misinformed, so that market outcomes will come to be dominated by those

Wald percentile (%)	Behavioral	Rational
Dynamics (VAR coeffs)	100	90.0
Volatility (variances)	96.0	24.2
Overall (all α)	100	79.8

Table 1: Comparison of Behavioural and Rational Expectations Models, indirectly re-estimated Parameters

with rational expectations.[25] Keynes (1930, 160) made the same point when he observed that "actions based on inaccurate anticipations will not long survive experiences of a contrary character, so that the facts will soon override anticipation except where they agree." In today's world, these observations would appear to have even greater force. For instance, Minford and Peel (2002) argue that rational expectations rests on the ability of competitive markets in information to process it efficiently: industries grow up to make these markets as efficient as possible – such as analysts, portfolio advisers, forecasters, hedge funds, and investment banks. The 'ordinary person' may not have literal rational expectations but is enabled to access sources that do have them. Under this argument, realistic but non-rational expectations are driven out by these sources, so that models based on them cannot fit the data behavior.[26]

6 Conclusion

We have surveyed recent theoretical literature comparing inflation targeting (IT) and price-level targeting (PT) as macroeconomic stabilization policies, focusing in particular on New Keynesian models and areas that have seen significant developments since Ambler's (2009) survey: the zero lower bound on nominal interest rates; financial frictions; and optimal monetary policy. The main conclusion reached by Ambler (2009) was that the ability of PT to improve the inflation-output volatility trade-off in New Keynesian models rests with the assumption of rational expectations. The recent literature suggests that things are somewhat more complicated than this. In particular, we highlight an important split in the literature: PT consistently outperforms IT when policy is described by simple Taylor-type rules, but results favouring PT which are derived from optimal policies – under either commitment or discretion – are

[25] Shiller (1978, p. 39) takes issue with this line of reasoning. He argues that if the economy takes a long time to converge on the rational expectations equilibrium, occassional changes in the structure of the economy may prevent a rational expectations equilibrium ever being attained. Shiller's argument is cited by Maddock and Carter (1982), who provide an excellent down-to-earth discussion of the macro implications of rational expectations.

[26] This argument relies on the assumption that the 'ordinary person' is not deliberately misled by sources with rational expectations who have an informational advantage. While examples of such behaviour are observed in the real world, these actions are unlikely to be systematic feature of the marketplace in competitive economies with strong property rights, because they would lead to punishment ex post either by withdrawal of custom or enforcement of the law.

fragile to the assumptions one makes about the economic environment faced by the policymaker, in the sense that they can be easily overturned with small deviations from the baseline New Keynesian model. We also argue, however, that Taylor-type rules are the case of most relevance for central banks in practice because, unlike optimal policies, rules are easy to implement and robust across alternative models. We thus emerge with a more precise version of Ambler's original conclusion that we view as relevant for real-world central banks: PT tends to improve macro stabilization in New Keynesian models with rational expectations when policymakers commit to Taylor-type rules.

Several of the extensions in the theoretical literature that we surveyed are initial attempts to address issues raised by the financial crisis that have led some to question the wisdom of central banks paying attention to PT, due to its reliance on rational expectations and New Keynesian models. A new finding that emerges from the recent literature is that PT is potentially very attractive in the context of the zero bound because, in conditions of deflationary recession, a price-level target induces expectations of higher than usual inflation. These expectations in turn induce negative real interest rates which stimulate economic activity out of the recession as well as putting an end to the deflation and the zero bound situation. This mechanism is much stronger than the equivalent mechanism under an inflation target and therefore raises the possibility of safely lowering trend inflation under a PT regime. The recent literature suggests that this dual mandate – a price-level target with a lower trend inflation rate – could bring substantial welfare gains, though the (strong) assumption of perfect credibility of PT is likely to be crucial for this result. There is also evidence from the recent literature that augmenting New Keynesian models with financial frictions leaves intact the potential welfare gains from PT, albeit that this literature is currently at an early stage.

Since the beneficial effects of PT appear to hang on the joint assumption that economic agents are rational and the economy New Keynesian, we devoted the second half of the survey to experimental and survey evidence on whether expectations are rational, and the applied macro literature on the empirical performance of New Keynesian models. In addition, we surveyed a more recent strand of applied literature that has formally tested New Keynesian models and alternative forms of expectation formation. Overall the evidence is not conclusive, but we showed that New Keynesian models are able to match a number of dynamic features in the data and that behavioral models of the macroeconomy do not appear to outperform those with rational expectations in formal statistical tests. Accordingly, we concluded that policymakers should continue to pay attention to PT in the future.

Bibliography

Adam, K. and Billi, R.M. 2007. Discretionary monetary policy and the zero lower bound on nominal interest rates. *Journal of Monetary Economics* 54(3), 728–52.

Adam, K. and Billi, R.M. 2006. Optimal monetary policy under commitment with a zero bound on interest rates. *Journal of Money, Credit and Banking* 38(7), 1877–1905.

Amano, R., Ambler, S., and Shukayev, M. 2012. Optimal Price-Level Drift under Commitment in the Canonical New Keynesian Model. *Canadian Journal of Economics* 45(3), 1023-36.

Amano, R., Engle-Warnick, J. and Shukayev, M. 2011. Price-level targeting and inflation expectations: experimental evidence. Bank of Canada Working Paper 2011-18.

Ambler, S. 2009. Price-level targeting and stabilisation policy: a survey. *Journal of Economic Surveys* 23(5), 974–97.

Ambler, S. 2007. The costs of inflation in New Keynesian models. *Bank of Canada Review* (Winter), 5–14.

Andolfatto, D., Hendry, S. and Moran, K. 2008. Are inflation expectations rational? *Journal of Monetary Economics* 55(2), 406–22.

Angrist, J. and Pischke, J.-S. 2010. The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. *Journal of Economic Perspectives* 24(2), 3–30.

ap Gwilym, R. 2010. Can behavioral finance models account for historical asset prices? *Economics Letters* 108(2), 187–89.

Bailey, M.J. 1956. The welfare cost of inflationary finance. *Journal of Political Economy* 64(2), 93–110.

Bailliu, J., Meh, C. and Zhang, Y. 2012. Macroprudential rules and monetary policy when financial frictions matter. Bank of Canada Working Paper 2012-6.

Barro, R. J. and Gordon, D.B. 1983. A positive theory of monetary policy in a natural rate model. *Journal of Political Economy* 91(4), 589–610.

Batini, N. and Yates, T. 2003. Hybrid inflation and price-level targeting. *Journal of Money, Credit and Banking* 35 (3), pp. 283-300.

Bernanke, B., Gertler, M. and Gilchrist, S. 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J.B. and Woodford, M. eds. Handbook of Macroeconomics. Amsterdam: North-Holland.

Bernatzi, S. and Thaler, R.H. 1995. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics* 110(1), 73–92.

Bils, M. and Klenow, P. 2004. Some evidence on the importance of sticky prices. *Journal of Political Economy* 112(5), 947–85.

Blake, A.P. and Kirsanova, T. 2012. Discretionary policy and multiple equilibria in LQ RE models. *Review of Economic Studies* 79 (4), 1309–39.

Blake, A.P., Kirsanova, T. and Yates, A. 2011. The gains from delegation revisited: price-level targeting, speed limit and interest rate smoothing policies. Bank of England Working Paper 415.

Blinder, A. S. 1997. "What Central Bankers Can Learn from Academics—and Vice-Versa," *Journal of Economic Perspectives* 11(2), 3–19.

Bloomfield, R. and Hales, J. 2002. Predicting the next step of a random walk: experimental evidence of regime-shifting beliefs. *Journal of Financial Economics* 65(3), 397–414.

Bonham, C.S. and Cohen, R.H. 2001. To aggregate, pool or neither: testing the rational expectations hypothesis using survey data. J*ournal of Business and Economic Statistics*, 19 (3), 278–291.

Canova, F. 2007. *Methods for applied macroeconomic research.* NJ: Princeton University Press.

Caplin, A., and Spulber, D. 1987. Menu Costs and the Neutrality of Money. *Quarterly Journal of Economics* 102, 703–725.

Carlson, J.A. and Parkin, M. 1975. Inflation expectations. *Economica* 42(1), 123–38.

Christiano, L.J., Eichenbaum, M.S. and Evans, C.L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1), 1–45.

Clarida, R., Galí, J. and Gertler, M .1999. The science of monetary policy: a new Keynesian perspective. *Journal of Economic Literature* 37(4), 1661–1707.

Coibion, O., Gorodnichenko, Y. and Wieland, J. 2012. The optimal inflation rate in New Keynesian models: should central banks raise their inflation targets in light of the zero lower bound? *Review of Economic Studies* 79, 1371–1406.

Coletti, D., Lalonde, R., Masson, P., Muir, D. and Snudden, S. 2012. Commodities and monetary policy: implications for inflation and price level targeting. Bank of Canada Working Paper 2012-16.

Coletti, D., Lalonde, R. and Muir, D. 2008. Inflation targeting and price-level-path targeting in the Global Economy Model: some open economy considerations. *IMF Staff Papers* 55(2), 326-338.

Costain, J. and Nakov, A. 2011. Price adjustments in a general model of state-dependent pricing. *Journal of Money, Credit and Banking* 43, 385–406.

Covas, F. and Zhang, Y. 2010. Price-level versus inflation targeting with financial market imperfections. *Canadian Journal of Economics* 43(4), 1302–32.

De Grauwe, P. 2010. Top-down versus bottom-up macroeconomics. *CESifo Economic Studies* 56(4), 465–97.

De Grauwe, P. and Grimaldi, M. 2006. *The exchange rate in a behavioral finance framework*. NJ: Princeton University Press.

Dhyne, E., Alvarez, L.J., Le Bihan, H., Veronese, G., Dias, D., Hoffmann, J., Jonker, N., Lunnemann, P., Rumler, F. and Vilmunen, J. 2006. Price setting in the Euro Area and the United States: some facts from individual consumer price data. *Journal of Economic Perspectives* 20, 171–192.

Dib, A., Mendicino, C. and Zhang, Y. 2013. Price-level targeting rules and financial shocks: the case of Canada. *Economic Modelling* 30(C), pp. 941–53.

Dib, A., Mendicino, C. and Zhang, Y. 2008. Price-level targeting in a small open economy with financial frictions: welfare analysis. Bank of Canada Working Paper No. 2008-40.

Dixon, H. and Le Bihan, H. 2012. Generalised Taylor and generalised Calvo wage and price setting: micro-evidence with macro implications. *Economic Journal* 122(560), 532–54.

Eggertsson, G.B. and Woodford, M. 2003. The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity* 1:2003, 139–211.

Evans, G.W. and Honkapohja, S.M.S. 2006. Monetary policy, expectations and commitment. *Scandinavian Journal of Economics* 108(1), 15–38.

Evans, G.W. and Honkapohja, S.M.S. 2001. *Learning and expectations in macroeconomics*. NJ: Princeton University Press.

Fair, R.C. 1983. Testing the rational expectations hypothesis in macroeconometric models. *Oxford Economic Papers* 45(2), 169–90.

Fernández-Villaverde, J. 2010. The econometrics of DSGE models. SERIEs 1(1), 3–49.

Fujiwara, I., Hirose, Y. and Shintani, M. 2011. Can news be a major source of aggregate fluctuations? *Journal of Money, Credit and Banking* 43, 1–29.

Gaspar, V., Smets, F. and Vestin, D. 2007. Is the time ripe for price level path stability? European Central Bank Working Paper 818.

Gerberding, C., Gerke, R. and Hammermann, F. 2012. Price-level targeting when there is price-level drift. *Journal of Macroeconomics* 34(3), 757–68.

Giannoni, M. 2010. Optimal interest rate rules in a forward-looking model, and inflation stabilization versus price-level stabilization. NBER Working Paper 15986.

Golosov, M. and Lucas, R.E. Jr. 2007. Menu costs and Phillips curves. *Econometrica* 115, 171–99.

Gourieroux, C. and Monfort, A. 1996. Simulation-based econometric methods. Oxford: Oxford University Press.

Hansen, L.P. and Heckman, J.J. 1996. The empirical foundations of calibration. *Journal of Economic Perspectives* 10(1), 87–104.

Holden, K., Peel, D.A., and Thompson, J.L. 1985. *Expectations – theory and evidence*. London: Macmillan.

Julliard, M. 2001. Dynare: a program for the simulation of rational expectations models. *Computing in Economics and Finance* 213, Society for Computational Economics.

Keynes, J.M. 1930. *A treatise on money*. Vols. I and II. NY: Harcourt Brace.

Khan, H.U. and Tsoukalas, J. 2012. The quantitative importance of news shocks in estimated DSGE models. *Journal of Money, Credit and Banking* 44(8), 1535–61.

Kydland, F.E. and Prescott, E.C. 1982. Time to build and aggregate fluctuations. *Econometrica* 50(6), 1345–70.

Kydland, F E. and Prescott, E.C. 1977. Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 85(3), 473–91.

Laxton, D. 2008. Getting to know the Global Economy Model and its philosophy. *IMF Staff Papers* 55(2), 213–42.

Le, V.P.M., Meenagh, D., Minford, P. and Wickens, M.R. 2011. How much nominal rigidity is there in the US economy? Testing a New Keynesian DSGE model using indirect inference. *Journal of Economic Dynamics and Control* 35(12), 2078–2104.

Levin, A., López-Salido, D., Nelson, E., Yun, T., 2010. Limitations on the effectiveness of forward guidance at zero lower bound. *International Journal of Central Banking* 6, 143-89.

Levine , P., Pearlman, J., Perendia, G. and Yang, B. 2012. Endogenous persistence in an estimated DSGE model under imperfect information. *Economic Journal* 122, 1287–1312.

Lim, G.C. and McNelis, P.D. 2008. *Computational macroeconomics for the open economy*. MA: MIT Press.

Liu, C. and Minford, P. 2012. Comparing behavioural and rational expectations for the US post-war economy. CEPR Discussion Paper 9132.

Lucas, R.E. Jr. 1976. Econometric policy evaluation: a critique. *Carnegie-Rochester Series on Public Policy* 1(1), 19–46.

Mackowiak, B. and M. Wiederholt (2011). Rational inattention and macroeconomic dynamics. ECB Working Paper No. 1331.

Maddock, R. and Carter, M. 1982. A child's guide to rational expectations. *Journal of Economic Literature* 20(1), 39–51.

Mankiw, N.G. and Reis, R., 2002. Sticky infomation versus sticky prices: a proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117(4), 1295–1328.

Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement about inflation expectations. *NBER Macroeconomics Annual* 18 (1), 209–248.

Masson, P.R. and Shukayev, M.D. 2012. Are bygones not bygones? Modeling price-level targeting with an escape clause and lessons from the gold standard. *Journal of Macroeconomics* 33(2), 162-75.

Meenagh, D., Minford, P. and Peel, D.A. 2007. Simulating stock returns under switching regimes – a new test of market efficiency. *Economics Letters* 94(2), 235–39.

Milani, F. (Forthcoming). The modeling of expectations in empirical DSGE models: a survey. Fortcoming in *Advances in Econometrics* 28, p. 1–36.

Minford, P. and Peel, D.A. 2002. *Advanced macroeconomics: a primer* (2nd Ed.), London: Blackwell.

Murchison, S. 2010. Price-level targeting and relative price shocks. *Bank of Canada Review* (Summer 2010), 11–21.

Muth, J.F. 1961. Rational expectations and the theory of price movements. *Econometrica* 29(3), 315–35.

Nakamura, E. and Steinsson, J. 2008. Five facts about prices: a reevaluation of menu cost models. *Quarterly Journal of Economics* 123(4), 1415–1464.

Nakov, A. 2008. Optimal and simple monetary policy rules with zero floor on the nominal interest rate. *International Journal of Central Banking* 4(2), 73–127.

Nessén, M. and Vestin, D. 2005. Average inflation targeting. *Journal of Money, Credit and Banking* 37(5), 837–63.

Rogoff, K. 1985. The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100(4), 1169–89.

Röisland, O. 2006. Inflation inertia and the optimal hybrid inflation/price-level target. *Journal of Money, Credit and Banking* 38(8), 2247–51.

Schmitt-Grohé, S., Uribe, M., 2012. What's news in business cycles?, *Econometrica* 80, 2733-2764.

Shiller, R.J. 1978. Rational expectations and the dynamic structure of macroeconomic models: a critical review. *Journal of Monetary Economics* 4(1), 1–44.

Smets, F. and Wouters, R. 2007. Shocks and frictions in US business cycles: a Bayesian DSGE approach. *American Economic Review* 97(3), 586–606.

Slobodyan, S. and Wouters, R. 2012. Learning in a medium-scale DSGE model with expectations based on small forecasting models. *American Economic Journal: Macroeconomics* 4(2), 65–101.

Smith, A.A. Jr. 1993. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics* 8, 63–84.

Smith, V.L., Suchanek, G.L. and Williams, A.W. 1988. Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica* 56(5), 1119–51.

Steinsson, J. 2003. Optimal monetary policy in an economy with inflation persistence. J*ournal of Monetary Economics* 50(7), 1425–56.

Svensson, L.E.O. 1999. Price-level targeting versus inflation targeting: a free lunch? *Journal of Money, Credit and Banking* 31(3), 277–95.

Svensson, L.E.O. 1997. Inflation forecast targeting: implementing and monitoring inflation targets. *European Economic Review* 41(6),1111–46.

Taylor, J.B. 1999. The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the European Central Bank. *Journal of Monetary Economics* 43(3), 655- 679.

Taylor, J.B. 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Series on Public Policy* 39, 195–214.

Taylor, J.B. 1980. Aggregate dynamics and staggered contracts, *Journal of Political Economy* 88(1), 1-23.

Thomas, L.B. Jr. 1999. Survey measures of expected U.S. inflation. *Journal of Economic Perspectives* 13(4), 125–144.

Vestin, D. 2006. Price-level versus inflation targeting. *Journal of Monetary Economics* 53(7), 1361–76.

Walsh, C.E. 2011. The future of inflation targeting. *Economic Record* 87(Supplement), 23–36.

Walsh, C.E. 2010. Monetary theory and policy (3rd Ed.). MIT Press.

Walsh, C.E. 1995. Optimal contracts for central bankers. *American Economic Review* 85(1), 150–67.

Wickens, M.R. *Macroeconomic theory: a dynamic general equilibrium approach* (2nd ed.). NJ: Princeton University Press.

Woodford, M. 2003. Interest and prices: Foundations of a theory of monetary policy. NJ: Princeton University Press.

Yates, A. 2004. Monetary policy and the zero bound to interest rates: a review. *Journal of Economic Surveys* 18(3), 427–81.