Article:
Cornea-Madeira, Adriana, Hommes, Cars and Massaro, Domenico (2019) Behavioral Heterogeneity in U.S. Inflation Dynamics. Journal of Business and Economic Statistics. pp. 288-300. ISSN: 0735-0015

https://doi.org/10.1080/07350015.2017.1321548

Behavioral Heterogeneity in U.S. Inflation Dynamics

Adriana Cornea-Madeira*
Management School, University of York, UK,
and
Cars Hommes
CenDEF, University of Amsterdam, and Tinbergen Institute, The Netherlands,
and
Domenico Massaro
Complexity Lab in Economics (CLE), Department of Economics and Finance,
Università Cattolica del Sacro Cuore, Milano

March 13, 2017

Abstract

In this paper we develop and estimate a behavioral model of inflation dynamics with heterogeneous firms. In our stylized framework there are two groups of price setters, fundamentalists and random walk believers. Fundamentalists are forward-looking in the sense that they believe in a present-value relationship between inflation and real marginal costs, while random walk believers are backward-looking, using the simplest rule of thumb, naive expectations, to forecast inflation. Agents are allowed to switch between these different forecasting strategies conditional on their recent relative forecasting performance. We estimate the switching model using aggregate and survey data. Our results support behavioral heterogeneity and the significance of evolutionary learning mechanism. We show that there is substantial time variation in the weights of forward-looking and backward-looking behavior. Although on average the majority of firms use the simple backward-looking rule, the market has phases in which it is dominated by either the fundamentalists or the random walk believers.

Keywords: Phillips Curve, Heterogeneous Expectations, Evolutionary Selection

*We thank Peter Boswijk and João Madeira for comments. This paper was presented at the Society for Computational Economics (2012), Asian Econometric Society Meeting (2013), European Econometric/Economic Society Meeting (2013), Bocconi University (2013), New York Federal Reserve Board (2013), Royal Economic Society (2014), International Economic Association World Congress (2014). Financial support from the British Academy's PDF/2009/370 and from the EU 7^{th} framework collaborative project "*Monetary, Fiscal and Structural Policies with Heterogeneous Agents (POLHIA)*", grant no. 225408 and from the EU 7^{th} framework collaborative project "*Complexity Research Initiative for Systemic InstabilitieS (CRISIS)*", grant no. 288501 are gratefully acknowledged.

1 Introduction

"Any time seems to be the right time for reflections on the Phillips curve"
Robert Solow

New Keynesian macroeconomics has greatly contributed to our understanding of inflation dynamics by considering models with nominal rigidities and optimizing agents with rational expectations. Despite their great popularity, New Keynesian models have also been the object of severe criticisms. Among others, Rudd and Whelan (2005a,b) question the lack of inflation inertia in the standard forward-looking New Keynesian Phillips curve (NKPC). Addressing this critique has resulted in various "hybrid" variants of the NKPC featuring both forward and backward-looking components, which have been theoretically motivated in several ways (see e.g., Fuhrer and Moore (1995), Christiano et al. (2005) and Galí and Gertler (1999)). Nevertheless, researchers continue to pursue more satisfying approaches to address inflation persistence (see Romer (2011) and Madeira (2015)). Another important criticism is the reliance of standard New Keynesian models on a strict form of rational expectations (RE) (see among others Rudd and Whelan (2006), and Carriero (2008)).[1]

To address these criticisms, this paper relaxes the RE assumption and proposes a model of inflation dynamics characterized by behavioral heterogeneous expectations. In our model agents form subjective beliefs (forecasts) about future inflation and can decide to switch to a different belief (forecasting rule) based on their forecasting performance. Our stylized model includes two types of firms. The first type are *fundamentalists*, who believe in a present-value relationship between inflation and real marginal costs. The second type, *random walk believers*, assume instead that inflation follows a random walk and use the simplest backward-looking rule of thumb corresponding to naive expectations (i.e., their forecast coincides with the last available observation) to forecast future inflation. We choose this

[1] The assumption of RE in the formulation of inter-temporal optimization decisions has also been criticized by Hendry and Mizon (2010, 2014) in the presence of unanticipated structural breaks on the grounds that the law of iterated expectations needs not to hold when distributions shift, as integrals are taken over different weighted intervals. Castle et al. (2014) find evidence for such shifts when fitting the hybrid NKPC to U.S. inflation data and demonstrate that a potentially spurious outcome can arise when the NKPC is estimated under the assumption of RE.

specific set of forecasting rules in order to obtain a NKPC similar to the closed-form solution of hybrid models estimated in the literature (see for example Sbordone (2005) and Rudd and Whelan (2006)). Our specification is more general as it allows agents to switch between rules based on their forecasting performance. Standard hybrid models as well as the purely forward-looking and the backward-looking Phillips curves can be seen as special instances of our general specification. Our setup is similar in spirit to the framework discussed in Hachem and Wu (2014), where monopolistically competitive firms make decisions before price level realizations and therefore rely on inflation forecasts. Firms have heterogeneous forecasting rules, namely a rule which is consistent with central bank's announcements, and a rule consistent with a random walk, while beliefs are updated over time via social dynamics on the basis of relative performance.[2] The authors then focus on the effectiveness of central bank communication and indentify how such communication can be tailored to build endogenous credibility. We focus instead on the estimation of a behavioral model with micro-founded time variation in the weights of forward- and backward-looking firms.

The results of our analysis provide empirical evidence for behavioral heterogeneity in U.S. inflation dynamics and in survey data of professional forecasters. Moreover, the data support the hypothesis of an endogenous mechanism relating predictors choice to their forecasting performance. In fact, our results suggest that the degree of heterogeneity varies considerably over time, and that the economy can be dominated temporarily by either forward-looking or backward-looking behavior.

The main assumptions in our behavioral model, i.e., heterogeneous firms with subjective forecasts and endogenous switching between different forecasting regimes on the basis of prediction performances, stem from two empirical stylized facts.

First, there is an abundant literature documenting heterogeneity in inflation expectations. Noticeably, Branch (2004), Carroll (2003), Pfajfar and Santoro (2010), Madeira and Zafar (2015), Mankiw et al. (2003) provide empirical evidence in support of heterogeneous expectations using survey data on inflation expectations.[3] Consistently with this litera-

[2] Arifovic et al. (2012) consider a DSGE model in which agents' beliefs evolve through social learning dynamics and show that the Taylor Principle is not necessary for convergence to the minimum state variable solution under social learning.

[3] Heterogeneity in individual expectations has also been found in other settings. For example, Frankel

ture we also find evidence on expectations heterogeneity using survey data on professional inflation forecasts.

Second, when hybrid NK models are estimated, there is mixed evidence about the importance of the forward-looking term and the backward-looking term. For example, Galí and Gertler (1999), Sbordone (2005), Kurmann (2007) and Kleibergen and Mavroeidis (2009) have found that the forward-looking component is more important than the backward-looking component. On the other hand, Lindé (2005) finds that the backward-looking component is equally or more important, while Madeira (2014) finds that it is predominant. Fuhrer (1997) and Rudd and Whelan (2006) conclude that the forward looking component plays essentially no role in observed inflation dynamics. This mixed evidence could be explained by the fact that there are periods when the forward-looking behavior is predominant and periods when the backward-looking behavior is more important. In particular, Carroll (2003) and Mankiw et al. (2003) show that inflation expectations evolve over time in response to economic volatility. In addition, Zhang et al. (2008), Kim and Kim (2008), Castle et al. (2014) and Hall et al. (2012) find evidence for multiple structural breaks in the coefficients of the forward- and backward-looking terms.[4] Our results provide a behavioral micro-foundation for the structural breaks observed in the relative weight of forward-looking term in the NKPC. This is also consistent with the micro evidence of Frankel and Froot (1991), Bloomfield and Hales (2002), Branch (2004) and Assenza et al. (2011), among others, who show that agents' expectations evolve in time as a response to the past forecast errors of agents using both survey and experimental data.

Our findings have important implications for monetary policy. Standard policy recommendations based on determinacy under RE may not be a robust criterion for policy advices in the presence of heterogeneous expectations. In fact, recent papers have shown that multiple equilibria, periodic orbits and complex dynamics can arise in the presence of

and Froot (1990), Allen and Taylor (1990) and Ito (1990) find that financial experts use different forecasting strategies to predict exchange rates, while Hommes et al. (2005), Adam (2007), Pfajfar and Zakelj (2014) and Assenza et al. (2011) find evidence for heterogeneity in learning to forecast laboratory experiments with human subjects.

[4]See also Stock and Watson (2007) and Groen et al. (2013) who acknowledge the possibility of the changing time-series properties of inflation.

dynamic predictor selection, even if the model under RE has a unique stationary solution; see, e.g., Anufriev et al. (2013), Branch and McGough (2010), and De Grauwe (2011).

The paper is structured as follows. Section 2 derives a NKPC with heterogenous expectations and endogenous switching dynamics. Section 3 presents the estimation results and describes the fit of the model. Section 4 discusses the robustness of the empirical results to alternative forecasting models. Section 5 performs out-of-sample forecasting exercises. Section 6 provides additional evidence for the behavioral switching mechanism using survey data on inflation expectations, while Section 7 concludes.

2 The model

This section derives a NKPC with heterogeneous, potentially nonrational expectations and endogenous switching between forecasting strategies.

2.1 The NKPC with heterogeneous expectations

We consider a model with monopolistic competition, staggered price setting and heterogeneous firms. There is a continuum of firms producing differentiated goods indexed by $j \in [0, 1]$. Each firm produces one good and has a production technology that uses labor as the only factor of production. The demand curve for product j takes the form:

$$Y_t(j) = Y_t(P_t(j)/P_t)^{-\eta},$$

where η is the Dixit-Stiglitz elasticity of substitution among differentiated goods, Y_t is the aggregator function defined as $Y_t = [\int_0^1 Y_t(j)^{(\eta-1)/\eta} dj]^{\eta/(\eta-1)}$, and P_t is the aggregate price level defined as $P_t = [\int_0^1 P_t(j)^{1-\eta} dj]^{1/(1-\eta)}$. Nominal price rigidity *à la* Calvo is introduced by allowing, in every period, only a fraction $(1 - \omega)$ of the firms to set a new price. Given that each firm hires labor from the same integrated economy-wide labor market, the prices chosen by the firms that can re-optimize in each period will only differ because of their subjective forecasts. Firms that reset prices maximize expected discounted profits, which are given by

$$\max_{P_{i,t}(j)} E_t^i \sum_{s=0}^{\infty} \omega^s Q_{t,t+s} \left(\frac{P_{i,t}(j)}{P_{t+s}} - mc_{t+s} \right) \left(\frac{P_{i,t}(j)}{P_{t+s}} \right)^{-\eta} Y_{t+s},$$

where $Q_{t,t+s}$ is the stochastic discount factor, mc_t are real marginal costs of production and $P_{i,t}(j)$ denotes the price set by a firm with subjective expectations of type i (E_t^i), producing good j. When each firm producing a certain good $j \in [0, 1]$ has a different subjective expectation of type $i \in [0, 1]$, we can without loss of generality use a single index, say i, to denote expectations' type and produced good.[5] Therefore the term $P_{i,t}$ will denote the price of a firm with subjective expectations of type i producing good i. Defining $q_{i,t}^* \equiv P_{i,t}^*/P_t$, where $P_{i,t}^*$ is the profit-maximizing price, and log-linearizing the first order conditions of this problem around a zero inflation steady state leads to

$$\hat{q}_{i,t}^* = (1 - \omega\delta)\widehat{mc}_t + \omega\delta E_t^i(\hat{q}_{i,t+1}^* + \pi_{t+1}), \tag{1}$$

where δ is the time discount factor, $\pi_t \equiv \hat{p}_t - \hat{p}_{t-1}$ is the inflation rate, and hatted variables denote log-deviations from steady state.[6] The relative average price set by optimizing firms is given by $\hat{q}_t = \int_i \hat{q}_{i,t}^*$. Log-linearizing the aggregate price level equation yields

$$\pi_t = \frac{1 - \omega}{\omega}\hat{q}_t. \tag{2}$$

Under the assumption of a representative firm with rational expectations, Eqs. (1) and (2) can be used to derive the standard NKPC:

$$\pi_t = \delta E_t \pi_{t+1} + \gamma mc_t, \tag{3}$$

which relates inflation, π_t, to next period's expected inflation and to real marginal costs, mc_t, where $\gamma \equiv (1-\omega)(1-\delta\omega)\omega^{-1}$ and we omitted hats for notational simplicity. Deriving an equation for inflation similar to Eq. (3) is not entirely obvious when expectations are heterogeneous. Following Kurz et al. (2013), it is possible to aggregate the individual pricing rules in order to obtain an aggregate supply equation of the form

$$\pi_t = \delta \bar{E}_t \pi_{t+1} + \gamma mc_t + \xi_t, \tag{4}$$

where $\bar{E}_t = \int_i E_t^i$ denotes the average expectation of individuals and the term ξ_t is defined as $\xi_t \equiv (1-\omega)\delta \int_i (E_t^i q_{i,t+1}^* - E_t^i q_{t+1})$.[7] Eq. (4) shows that, in the presence of heterogeneous

[5] The model derivation in the presence of a discrete number of belief types, which is a particular case of the general specification derived in this section, is outlined in Appendix A.

[6] See Appendix A for a detailed derivation.

[7] See Appendix A for details.

expectations, inflation depends on real marginal costs, on the average forecasts of future inflation, and on an additional term ξ_t representing the difference between the average firms' forecast of individual prices $q^*_{i,t+1}$ and the average forecast of average price q_{t+1}. In the presence of heterogeneous agents, with possibly non-rational beliefs, there is no a-priori reason to believe that in every period the average forecast of individual prices will coincide with the average forecast of average price. Given that we have no data on such deviations, in our empirical analysis we will consider ξ_t as part of the error term and performs diagnostic checks on the properties of the residuals of our regression model.[8] This is in line with the ideas of, e.g., Kurz et al. (2013) and Diks and van der Weide (2005), who consider expectations heterogeneity as a natural source of randomness.

2.2 Evolutionary selection of expectations

We assume that agents form expectations by choosing from I different forecasting rules, and we denote by $E^i_t\pi_{t+1}$ the forecast of inflation by rule i. The fraction of individuals using the forecasting rule i at time t is denoted by $n_{i,t}$. Fractions are updated in every period according to an *evolutionary fitness* measure. At the beginning of every period t, agents compare the realized relative performances of the different strategies and the fractions $n_{i,t}$ evolve according to a discrete choice model with multinomial logit probabilities (see Manski and McFadden (1981) for details), that is

$$n_{i,t} = \frac{\exp(\beta U_{i,t-1})}{\sum_{i=1}^{I} \exp(\beta U_{i,t-1})}, \tag{5}$$

where $U_{i,t-1}$ is the realized fitness metric of predictor i at time $t-1$. The parameter $\beta \geq 0$ refers to the *intensity of choice* and reflects the sensitivity of the mass of agents to selecting the optimal prediction strategy. Brock and Hommes (1997) proposed this model for endogenous selection of expectation rules. We remark that Eq. (5) can also be derived

[8]Notice also that we cannot directly impose a structure on ξ_t since we will make assumptions about how agents forecast inflation but not about how agents forecast prices. From a behavioral point of view, forecasting prices is rather different than forecasting inflation (see e.g., Tuinstra and Wagener (2007)). In fact, while we will make specific assumptions on inflation expectations on the basis of observable statistical or theoretical properties of the inflation process, it is more difficult to model price expectations since in reality agents rarely collect information or read news about prices in levels.

from an optimisation problem under rational inattention (see Matějka and McKay (2015)). In the context of rational inattention the parameter β is inversely related to the *shadow cost of information*. The key feature of Eq. (5) is that strategies with higher fitness in the recent past attract more followers. The case $\beta = 0$ corresponds to the situation in which differences in fitness cannot be observed, so agents do not switch between strategies and all fractions are constant and equal to $1/I$. The case $\beta = \infty$ corresponds to the "neoclassical" limit in which the fitness can be observed perfectly and in every period all agents choose the best predictor.

2.3 A simple two-type example

In order to be able to compare our results with previous empirical works on the NKPC featuring a forward- and a backward-looking component, we assume that agents can choose between two forecasting rules to predict inflation, namely fundamental and random walk beliefs. The fundamental rule is based on a present-value description of the inflation process. When all agents have rational expectations, repeated application of Eq. (4) gives

$$\pi_t = \gamma \sum_{k=0}^{\infty} \delta^k E_t mc_{t+k}. \tag{6}$$

We refer to (6) as the *fundamental inflation*. Fundamentalists use expression (6) to forecast future inflation. In particular, leading (6) one-period ahead we get

$$\pi_{t+1} = \gamma \sum_{k=1}^{\infty} \delta^{k-1} E^f_{t+1} mc_{t+k}, \tag{7}$$

where E^f denotes fundamentalists forecast. Applying the expectation operator E^f_t on both sides we get

$$E^f_t \pi_{t+1} = \gamma \sum_{k=1}^{\infty} \delta^{k-1} E^f_t mc_{t+k}. \tag{8}$$

In deriving Eq. (8) we made use of the law of iterated expectations at the *individual* level. This is a reasonable and intuitive assumption which is standard in the learning literature; see, e.g., Evans and Honkapohja (2001) and Branch and McGough (2009).[9]

[9] We justify the fact that the law of iterated expectations holds at the individual level in the presence of evolutionary switching by appealing to the learning literature which models the selection of forecasting

From a behavioral point of view, fundamentalists can be considered as agents who believe in RE and use the closed form solution of the model to forecast the inflation path. There is, however, an important difference between fundamental expectations and RE. Fundamental expectations are *not* model-consistent because they do not take into account the presence of non-rational agents. Intuitively, in a world with heterogeneous firms, model-consistent expectations would require agents to collect an incredible amount of information about the economy, including details about the beliefs of other agents in the market, in order to derive the objective probability distribution of aggregate variables. More realistically, firms in our framework only have knowledge of their objectives and of the constraints that they face, and therefore they do not have a complete model of determination of aggregate variables.

In order to characterize the fundamental forecast (8) we use the VAR methodology of Campbell and Shiller (1987). Assuming that the forcing variable mc_t is the first variable in the multivariate VAR

$$Z_t = AZ_{t-1} + \epsilon_t, \tag{9}$$

we can rewrite the sum of discounted future expectations of marginal costs (8) as

$$E_t^f \pi_{t+1} = \gamma \sum_{k=1}^{\infty} \delta^{k-1} E_t^f mc_{t+k} = \gamma e_1' (I - \delta \hat{A}_{(t)})^{-1} \hat{A}_{(t)} Z_t,$$

where e_1' is a suitably defined unit vector and $\hat{A}_{(t)}$ is the recursive estimate of matrix A using the information set available at time t.[10] In fact, fundamentalist agents do not have perfect foresight into the future and, in every period t, they estimate a VAR model to produce their forecast using the information set available at time t. The coefficient matrix is then updated in every period as new information becomes available.

The second rule, which we call random walk belief or naive expectation, takes advantage of inflation persistence and uses a simple backward-looking forecasting strategy:

$$E_t^n \pi_{t+1} = \pi_{t-1}, \tag{10}$$

rules as a distinct statistical problem. Thus agents choose a forecasting model and then use that model to solve for their optimal plan in the anticipated utility sense, as in Kreps (1998) and Sargent (1999).

[10] Technically, because the discounted sum of real marginal costs starts at $k = 1$, we measure it using $(I - \delta \hat{A}_{(t)})^{-1} \hat{A}_{(t)} Z_t$ instead of $(I - \delta \hat{A}_{(t)})^{-1} Z_t$.

where E^n denotes the naive expectation operator. Notice that, although being an extremely simple rule, the naive forecasting strategy is optimal when the stochastic process is a random walk; hence for a near unit root process, as in the case of inflation, random walk expectations are close to optimal.

At this point a discussion on the choice of the set of forecasting rules is warranted. Admittedly, the fundamentalist prediction rule is built on the basis of a simple, stylized NKPC, while there are of course more complex versions of the NKPC including e.g., indexation and an autocorrelated price mark-up process (see Smets and Wouters (2007) for an example). Moreover, given the volatility of the inflation process, one may postulate a slightly more sophisticated naive forecasting strategy, such as an average over time of past inflation (see Section 4). Nevertheless, the specific choice of the set of forecasting rules, namely the fundamental rule in Eq. (8) and random walk beliefs in Eq. (10), will enable us to compare the outcome of our analysis with the results of previous empirical works based on the hybrid Phillips curve specification. In fact, fundamental expectations account for the forward-looking component in the estimated closed-form solution of the hybrid NKPC. We estimate the discounted sum of expected marginal costs using the VAR methodology as in Sbordone (2005) and Rudd and Whelan (2006) among others, with the only difference that, consistently with the behavioral interpretation of our model, the matrix of coefficients in the VAR model is not estimated using observations over the full sample (i.e., information which is not available to agents at the moment of their forecast) but it is recursively updated as new observations become available. In Section 4 we also consider Bayesian estimation of matrix A as a robustness check. The backward-looking component introduced in different ways in hybrid RE models is accounted for by the expectations of naive firms.

The main difference between traditional hybrid specifications of the NKPC and our model is the fact that the weights assigned to forward-looking and backward-looking component are endogenously varying over time. We assume that agents can switch between the two predictors based on recent forecasting performance.[11] Defining the absolute forecast

[11] In principle the switching metric should depend on the degree of price stickiness, i.e., the stickier the prices the longer the horizon the firms should take into account. We leave this issue for future research and, in what follows, we consider the average forecast error over the previous K periods as switching metric.

error in the previous K periods as

$$FE_t^i = \sum_{k=1}^{K} |E_{t-k}^i \pi_{t-k+1} - \pi_{t-k+1}|,$$

with $i = f, n$, we can then define the evolutionary fitness measure as[12]

$$U_{i,t} = -\frac{FE_t^i}{\sum_{i=1}^{I} FE_t^i}. \tag{11}$$

The evolution of the weights of different heuristics is then given by Eq. (5). Denoting the fraction of fundamentalists as $n_{f,t}$ we can summarize the full model as

$$\pi_t = \delta(n_{f,t} E_t^f \pi_{t+1} + (1 - n_{f,t}) E_t^n \pi_{t+1}) + \gamma mc_t + u_t, \tag{12}$$

where u_t is a composite error term including the component ξ_t and potential errors due to measurement or linearization, and

$$\begin{aligned}
E_t^f \pi_{t+1} &= \gamma e_1' (I - \delta \hat{A}_{(t)})^{-1} \hat{A}_{(t)} Z_t \\
E_t^n \pi_{t+1} &= \pi_{t-1} \\
n_{f,t} &= \frac{1}{1 + \exp\left(\beta \left(\frac{FE_{t-1}^f - FE_{t-1}^n}{FE_{t-1}^f + FE_{t-1}^n}\right)\right)} \\
FE_{t-1}^i &= \sum_{k=1}^{K} |E_{t-k-1}^i \pi_{t-k} - \pi_{t-k}|, \qquad \text{with } \ i = f, n \ .
\end{aligned}$$

3 Estimation results

This section describes the data and methodology used to estimate the nonlinear switching model derived in the previous section.

3.1 Data description

We use quarterly U.S. data on the inflation rate, the output gap, the labor share of income, and consumption-output ratio, from 1968:Q4 to 2015:Q2. The reason for the starting date

[12]The estimation results are robust to alternative specifications of the fitness measure. We chose relative absolute forecast error for numerical convenience, since it restricts the support of the fitness measure to the interval $[-1, 0]$.

of our sample is that observations on inflation expectations from the Survey of Professional Forecasters, to which we compare our model-implied measure of expectations in Section 6, are available from 1968:Q4 onwards.[13] Inflation is measured as log difference of the GDP price index. Output gap is measured as quadratically detrended log real GDP. We use labor share of income and detrended consumption-output ratio time series for nonfarm business sector in the construction of the VAR model (9). A more detailed description of data sources and variables definition is given in Appendix B.

3.2 The fit of the model

In this section we discuss the empirical implementation of model (12). In the "baseline" specification (the one used in the results reported below) we use the labor share of income as a proxy for real marginal costs (see, e.g., Galí and Gertler (1999), Woodford (2001) and Sbordone (2002) among others), and in Appendix F we consider the output gap as a measure of inflationary pressure.

Baseline VAR specification

The first step concerns the choice of the baseline VAR specification to estimate the matrix A at each point in time t, needed to construct the forecasts of fundamentalists,

$$E_t^f \pi_{t+1} = \gamma e_1'(I - \delta \hat{A}_{(t)})^{-1} \hat{A}_{(t)} Z_t.$$

We choose a four-lag bivariate VAR in the labor share of income (lsi_t) and the output gap (y_t) as our baseline specification. The optimal number of lags was chosen on the basis of the comparison of standard information criteria for the VAR estimated over the full sample. The Akaike information criterion, the Hannan-Quinn information criterion and the Final Prediction Error criterion selected a number of lags equal to 4, while the Bayesian information criterion opted for 2 lags. Given that 3 out of 4 criteria indicated 4 as the optimal lag length, we opted for a four lag baseline VAR. In Section 4 we perform robustness checks to alternative specifications of the forecasting VAR as well as Bayesian

[13] We repeated our estimation exercise using samples with different starting dates and results are not qualitatively different.

estimation of the fundamentalist forecast. Denoting by Y_t the vector of dependent variables, $Y_t = [lsi_t, y_t]'$, the vector Z_t is therefore defined as $Z_t = [Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3}]'$. Although being parsimonious, our VAR specification captures about 94% of the labor share of income volatility and the F test reports no autocorrelation in the residuals up to the 20^{th} lag. We initialize the VAR coefficients' estimates using all available pre-sample data from 1947:Q2. The matrix $\hat{A}_{(t)}$ is then updated in every period as new information becomes available.

NLS estimation

As standard in empirical works on the NKPC, we fix the discount factor $\delta = 0.99$, and we select a number of $K = 4$ lags for measuring past performance.[14] That is, if fundamentalists (naive) have a more accurate inflation forecast over the past year, more firms will follow the fundamentalist (naive) expectation formation rule. Model (12) is then estimated using non-linear least squares (NLS). Table 1 presents the results and diagnostic checks of the residuals are reported in Appendix C.

Table 1: NLS estimates of model (12)

Parameter	β	γ
Estimate	5.040***	0.001***
Std. error	1.641	0.0002
R^2 from Inflation Equation	0.82	

Notes: Standard errors are computed using White's heteroskedasticity-consistent covariance matrix estimator (HCCME). *, **, *** denote significance at the 10%, 5%, and 1% level.

All coefficients have the correct sign and are significant at the 1% level. The estimate of coefficient γ is in line with most estimates in the literature, while the positive sign and the significance of the intensity of choice parameter β implies that agents switch towards

[14] From a behavioral point of view it seems a sensible choice to pick $K = 4$ for quarterly data, meaning that the fitness measure takes an average of the forecast errors over the past year. Experimentation with different values of K shows that our results are robust to the choice of the number of lags in the performance measure.

the better performing forecasting rule, based on its past performance.[15] The time series of inflation predicted by model (12) for the estimated values of β and γ is plotted in Fig. 1, as dashed line, versus the actual series (solid line).[16] Overall the predicted inflation path

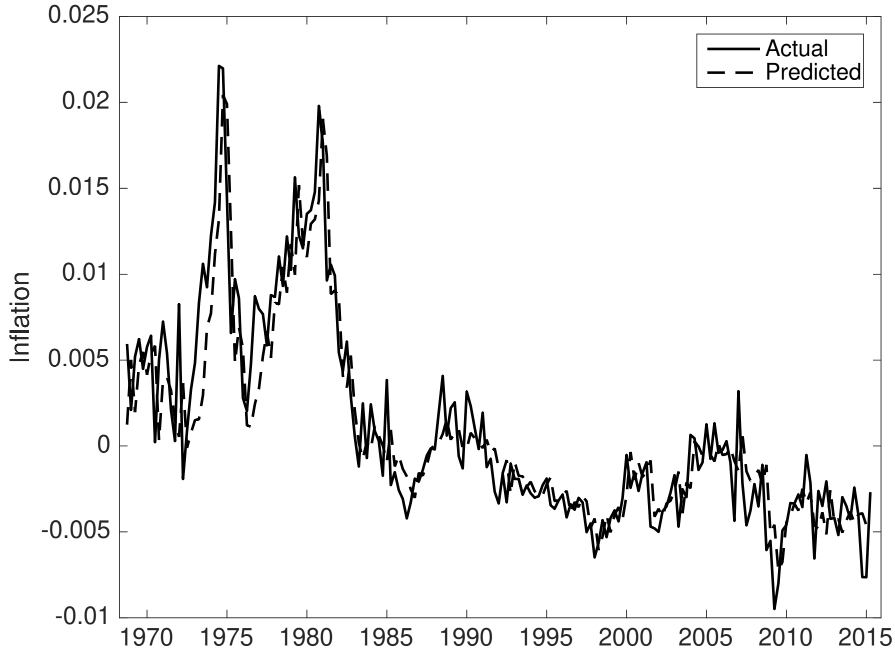

Figure 1: Actual versus predicted inflation

tracks the behavior of actual inflation quite well (the R^2 from inflation equation (12) is about 0.82, see Table 1).

Our results are, in some respects, similar to findings obtained in previous empirical works. In particular, Galí and Gertler (1999) and Sbordone (2005) find that models derived from the assumption of heterogeneous price setting behavior are capable of fitting the level of inflation quite well. However, Rudd and Whelan (2005a) and Rudd and Whelan (2006) show that this good fit reflects the substantial role that these models still allow for lagged inflation, and that forward-looking components play no discernable empirical role in determining inflation.

[15]The order of magnitude of β is more difficult to interpret as it is conditional on the functional form of the performance measure U.

[16]The series are in deviation from the mean.

Our NKPC specification allows for time-varying weights assigned to fundamentalists and naive price setters. Having estimated model (12), we are now ready to assess the relative importance over time of forward-looking versus backward-looking components in inflation dynamics. Table 2 displays descriptive statistics of the weight of the forward-looking component n_f.

Table 2: Descriptive statistics of the weight n_f

Mean	Median	Max	Min	St. Dev.	Skewness	Kurtosis	1^{st} order AC
0.353	0.276	0.924	0.019	0.282	0.418	1.720	0.887

On average, the majority of agents use the simple backward-looking rule (with average fraction $1 - 0.35 = 0.65$). However, the spread between the minimum and the maximum indicates that the market can be dominated by either forward-looking or backward-looking agents. Moreover the autocorrelation of the series n_f, about 0.89, indicates that agents do not change their strategy quickly, suggesting a relatively high degree of inertia in the updating process.

Fig. 2 shows the time series of the fraction of fundamentalists, i.e., the forward-looking component in our NKPC specification, the time series of the distance of actual inflation from the fundamental and the random walk forecasts, and a scatter plot of the fraction of fundamentalists against the relative forecast error of the naive rule. It is clear from this figure that the fraction of fundamentalists varies considerably over time with periods in which it is close to 0.5 and other phases in which it is close to either one of the extremes 0 or 1. For example, immediately after the oil crisis of 1973, the proportion of fundamentalists drops almost to 0. Soon after the difference between inflation and fundamental value reaches its peak in 1974:Q4, the estimated weight of the forward-looking component shoots back up to about 0.6. During the second oil crisis, inflation was far above the fundamental, causing more and more agents to adopt a simple backward-looking rule to forecast inflation. Our findings suggest that, in reaction to large shocks pushing inflation away from the fundamental, a large share of agents adopt random walk beliefs causing self-fulfilling high inflation persistence.

This result is in line with the analysis of Branch and Evans (2016) showing that innova-

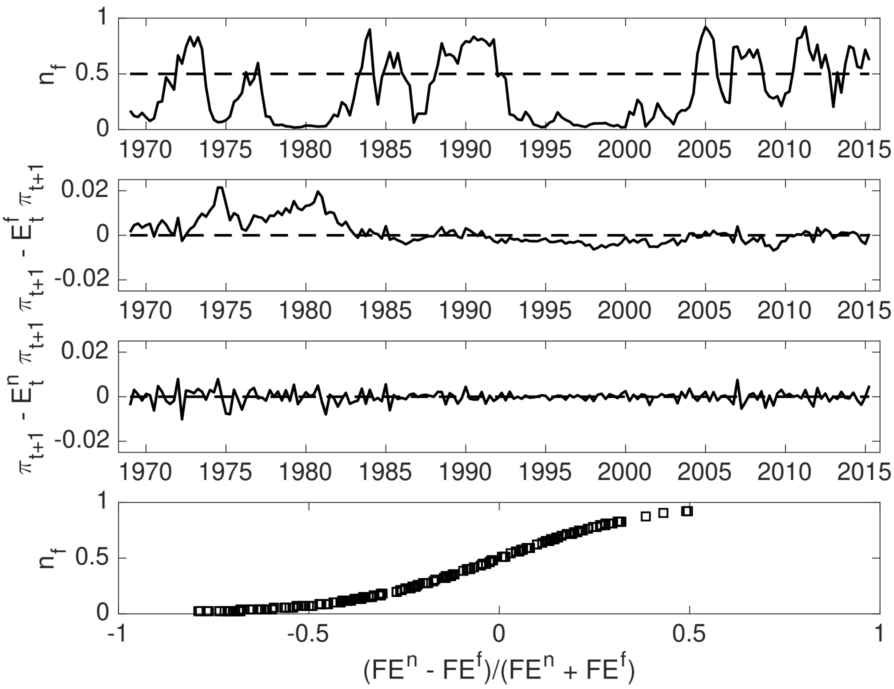

Figure 2: **First (top) panel:** Time series of the fraction of fundamentalists $n_{f,t}$. **Second panel:** Distance between actual inflation and fundamental forecast. **Third panel:** Distance between actual inflation and naive forecast. **Fourth (bottom) panel:** Scatter plot of the weight $n_{f,t}$ versus the relative forecast error of the naive rule.

tions to inflation can lead agents adaptively learning in the economy to temporarily believe that inflation follows a random walk. Fig. 2 shows that the fundamentalists dominated the economy in the late 1980s and early 1990s, while in the late stage of the Great Moderation (from 1992 until 2004), inflation stayed continuously well below the fundamental, causing the weight of random walk believers to increase. This result is consistent with the findings of Stock and Watson (2007), and Madeira and Zafar (2015) who show that the statistical capability of the last observed inflation for the next year inflation forecasts increased substantially in the period 1990-2005. We observe from Fig. 2 that in the aftermath of the global financial crisis of 2007/8 n_f increases, reaching peaks of about 0.9. Moreover, the heterogeneity of inflation expectations has increased since 2006, a finding consistent with the results of Madeira and Zafar (2015). This could be explained perhaps by the greater uncertainty due to the economic crisis. The bottom panel of Fig. 2 presents a scatter plot of the relative forecast error of the random walk rule, $(FE^n - FE^f)/(FE^n + FE^f)$, versus the fraction of fundamentalist agents, n_f. Due to the positive estimated value of β, this line slopes upwards, such that a more accurate fundamentalist forecast results in a higher weight n_f. The S shape is induced by the logit function in Eq. (5).

The analysis conducted in this section shows that the evolutionary learning model fits the data well. The positive sign and the significance of the intensity of choice parameter, β, imply that the endogenous mechanism that relates predictors choice to past performance is supported by the data. We also find that the ability of the discounted sum of expected future marginal costs values to predict the empirical inflation process varies considerably over time. In fact the spread between the minimum and the maximum value of n_f, i.e., the fraction of fundamentalists, shows that the economy can be dominated by either forward-looking or backward-looking behavior. Moreover, even though the market is, on average, dominated by agents using a simple heuristic to predict inflation, fundamentalists, or forward-looking components, still have a significant impact on inflation dynamics.

3.3 Specification tests and model selection

In order to assess the validity of our baseline heuristic switching model, which will be denoted by HSM, we test it against four alternative specifications: a model with heterogeneous agents and exogenous estimated fixed weights ($n_{f,t} \equiv \widehat{n}_f$), which is similar to the hybrid NKPC estimated by Rudd and Whelan (2006) and Sbordone (2005), and it is denoted by M_1; a model with homogeneous fundamentalists agents ($n_{f,t} \equiv 1$), which resembles the RE closed form solution of the standard NKPC without backward-looking component, denoted by M_2;[17] a model with homogeneous naive agents ($n_{f,t} \equiv 0$), which

[17] In fact, in the presence of homogeneous firms we have that $\xi_t = 0$ and, substituting the fundamental forecast in Eq. (4), we get $\pi_t = \delta\gamma \sum_{k=1}^{\infty} \delta^{k-1} E_t^f mc_{t+k} + \gamma mc_t = \gamma \sum_{k=0}^{\infty} \delta^k E_t^f mc_{t+k}$ which corresponds to the inflation path implied by the RE closed form solution in Eq. (6), when the discounted sums of current and future expected marginal costs are estimated in the same way. The difference between the model with only fundamentalists and the standard model under RE is the fact that typically in the latter the matrix of coefficients in the VAR model (9) is estimated using the full sample, i.e., $\hat{A}_{(T)}$, while consistently with our behavioral model, fundamentalists only use available information in each period implying that the matrix of coefficients A is estimated period by period, i.e., $\hat{A}_{(t)}$. The outcome of the test reported in Table 3 below

recalls the backward-looking Phillips curve and it is denoted by M_3; a static model with heterogeneous agents in which we let $\beta = 0$ ($n_{f,t} \equiv 0.5$), which is similar to the model of Fuhrer and Moore (1995), denoted by M_4. Given that, with the exception of model M_4 which obtains by setting $\beta = 0$, the competing models are nonnested, we will use nonnested hypothesis testing procedures. In particular, we construct the P test for the adequacy of our nonlinear specification with endogenous switching in explaining inflation dynamics (null hypothesis) against the alternative specifications mentioned above. Nonnested hypotheses tests are appropriate when rival hypotheses are advanced for the explanation of the same economic phenomenon. We will follow the procedure described in Davidson and MacKinnon (1981) and Davidson and MacKinnon (2009) and compute a heteroskedasticity-robust P test of HSM against the alternatives M_1, M_2, M_3, and M_4. We report the results of the test in Table 3 and refer the reader to Davidson and MacKinnon (2009), p. 284 and p. 669, for details on the construction of the heteroskedasticity-robust test. The first two rows of

Table 3: Paired nonnested hypotheses tests and Bayesian Information Criteria

	HSM	M_1	M_2	M_3	M_4
HSM vs. M_c	–	0.175	0.450	0.180	0.959
M_c vs. HSM	–	0.020	0.000	0.001	0.034
BIC	**-1690**	−1681	−1584	−1394	−1675

Notes: The first row reports p-values from P tests of HSM model against each model M_c ($c \in \{1, 2, 3, 4\}$). The second row reports p-values from P tests of each model M_c ($c \in \{1, 2, 3, 4\}$) against HSM model. The third row reports the BIC for all models (best model shown in bold).

Table 3 report the results of paired nonnested tests in which we test the benchmark HSM against M_1, M_2, M_3, and M_4 (first line of the table), and we test each M_1, M_2, M_3, and M_4 model against the HSM model (second line of the table).[18] The first row of Table 3 shows

does not change if we substitute model M_2 with the actual RE closed form solution. Moreover, in Section 4 we estimate our behavioral model by giving the fundamentalists the same "informational advantage" as in standard models with RE, hence using the full sample to estimate the VAR. The qualitative results in Table 3 are not altered when considering a behavioral model in which the fundamentalists have an informational advantage as in RE models.

[18]For completeness, we also compared the HSM model to the nested static model M_4 using a likelihood ratio test. We rejected the null of a restricted static model at the 1% level.

that we never reject, with 95% confidence level, our baseline switching model (HSM) when tested against all alternative models (M_1, M_2, M_3, and M_4). In other words, there is no statistically significant evidence of departure from the null hypothesis, i.e., adequacy of the nonlinear switching model, in the direction of these alternative explanations of inflation dynamics. On the contrary, the second row of Table 3 shows that we reject each of the alternative models when tested against the switching model.[19]

The results of the nonnested hypotheses tests show that the data support our model of inflation dynamics and provide evidence in favor of the switching model when tested against alternative models of the inflation process. However, since nonnested hypotheses tests are designed as specification tests, we complete the analysis by reporting the Bayesian Information Criterion (BIC) for model selection in the last row of Table 3. The BIC chooses the baseline switching model as the best model among all competing models, confirming the results of the nonnested hypotheses tests.

In Appendix D we test our baseline model against a set of alternative models in which a representative agent forms forecasts using a constant gain algorithm.[20] We consider three different specifications for the perceived law of motions used by the agents to forecast future inflation (see Appendix D for details). Both the nonnested hypotheses tests and the comparison of BICs are in favor of the switching model, a result that is in line with the findings reported in this section.

4 Robustness analysis

In this section we address the issue of how sensitive our results are to alternative specifications of the forecasting rules of both fundamentalists and random walk believers. In order to investigate the robustness of our results to alternative specifications of the fundamentalist VAR forecasting model, we extend our baseline specification by adding lagged

[19] For the sake of completeness, we also test the joint significance of the alternative models against our benchmark nonlinear switching model. The resulting F-statistic is 4.828, which leads to the acceptance of the null hypothesis of joint insignificance of alternative models M_1, M_2, M_3, and M_4 against HSM (p-value $\chi^2(4) = 0.305$).

[20] We thank an anonymous referee for this suggestion.

inflation (π_{t-1}) and consumption-output ratio (c_t/y_t).[21] These variables have been used in the VAR specifications considered by Rudd and Whelan (2005a) and Sbordone (2002). For each of the VARs considered in this section, the number of lags has been chosen optimally by comparing standard information criteria as in Section 3.2 and picking the number of lags selected by the majority of the criteria.

Table 4: Estimation results using alternative VARs for labor share of income

	VAR specification					
	Recursive $\hat{A}_{(t)}$ (1–3)			Full sample $\hat{A}_{(T)}$ (4–6)		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\begin{bmatrix} lsi_t \\ y_t \end{bmatrix}$	$\begin{bmatrix} lsi_t \\ y_t \\ \pi_{t-1} \end{bmatrix}$	$\begin{bmatrix} lsi_t \\ y_t \\ \pi_{t-1} \\ c_t/y_t \end{bmatrix}$	$\begin{bmatrix} lsi_t \\ y_t \end{bmatrix}$	$\begin{bmatrix} lsi_t \\ y_t \\ \pi_{t-1} \end{bmatrix}$	$\begin{bmatrix} lsi_t \\ y_t \\ \pi_{t-1} \\ c_t/y_t \end{bmatrix}$
β	5.041***	5.611***	5.352***	7.030***	9.265***	8.957***
	(1.641)	(1.753)	(1.655)	(2.525)	(1.802)	(3.175)
γ	0.001***	0.004***	0.003***	0.004***	0.013***	0.006***
	(0.0002)	(0.0008)	(0.0006)	(0.0005)	(0.0013)	(0.0007)
	R^2 from Inflation Equation					
	0.82	0.82	0.82	0.81	0.79	0.81

Notes: $lsi_t \equiv$ labor share, $y_t \equiv$ output gap, $\pi_{t-1} \equiv$ (past) inflation, $c_t/y_t \equiv$ detrended consumption-output ratio. Lag length (l_i) in VAR specifications ($i = 1, \ldots, 3$): $l_1 = 4$, $l_2 = l_3 = 2$. Standard errors are computed using White's heteroskedasticity-consistent covariance matrix estimator (HCCME). *, **, *** denote significance at the 10%, 5%, and 1% level. Columns 1–3 report estimates for coefficient matrix $\hat{A}_{(t)}$ updated in every period, while columns 4–6 report estimates for coefficient matrix $\hat{A}_{(T)}$, estimated over the full sample.

Columns 1–3 of Table 4 report estimation results from alternative VAR forecasting models for the labor share of income, with coefficient matrix $\hat{A}_{(t)}$ updated by fundamen-

[21] We use π_{t-1} in the construction of the VAR to be consistent with the information set of fundamentalist firms in the model. In fact, as standard in learning models, current values of endogenous variables are not observable at time t because they depend on the heterogeneous beliefs in the economy which are not known to the individual firm.

talists in every period as new information becomes available. As shown in Table 4, the point estimates and significance of the coefficients β and γ do not change substantially. As a further robustness check, in columns 4–6 we repeat the exercise and estimate the VAR coefficient matrix over the full sample, i.e., $\hat{A}_{(T)}$, in the spirit of standard RE models. The results support behavioural heterogeneity and switching behavior even in the presence of such "informational advantage" for fundamentalists.

We also estimate our behavioral model using different BVARs in the estimation of the fundamentalist forecast. We consider two different BVAR specifications: a BVAR with two variables, namely lsi_t and y_t, with four lags, and a BVAR with four variables, namely lsi_t, y_t, π_{t-1} and c_t/y_t with four lags. For each BVAR we consider three different prior specifications, i.e., a Sims-Zha Normal Wishart prior, a Minnesota prior, and a simpler Normal Wishart prior (see Appendix E for details on the specification of the priors' hyper-parameters). The results are reported in Table E.2, Appendix E. The estimates of both coefficients β and γ are positive and statistically significant and the results are robust to variations in the priors' hyper-parameters.

As an additional robustness check, we estimate the switching model assuming that the naive predictor is given by the average of the past four lags of inflation along the lines of Atkeson and Ohanian (2001). The naive forecasting rule thus reads as follows:

$$E_t^n \pi_{t+1} = \frac{1}{4} \sum_{k=1}^{4} \pi_{t-k}.$$

This specification is justified on the grounds that quarterly inflation data are volatile and therefore it is plausible that even naive agents use some average over time of past inflation. The results in Table E.3, Appendix E report once again positive and significant estimates of the parameters β and γ.

Overall, the robustness checks performed in this section suggest that our analysis is robust to different VAR forecasting models for the driving variable in the inflation process and to alternative naive forecasting rules. In Appendix F we investigate the sensitivity of our results to alternative measures of marginal costs. We re-estimate our model using an output gap measure as the driving variable in the inflation process and we find significant evidence for the behavioral switching mechanism.

5 Out-of-sample forecast

In this section we compare the point forecast accuracies of our heuristic switching baseline model (HSM) and models M_1 ($n_{f,t} \equiv \hat{n}_f$), M_2 ($n_{f,t} \equiv 1$), M_3 ($n_{f,t} \equiv 0$) as described in Section 3.3 to a benchmark model for inflation forecasts, namely the Unobserved Components model with Stochastic Volatility (UC-SV) of Stock and Watson (2007) (see Appendix G for details). In the following we construct iterated point forecasts from these models based on the recursive window scheme (the estimates are updated at each forecast origin using all available information).[22] In forming the first forecast, we estimate the models using data from 1968:Q4 (consistently with the analysis in the previous sections) to 1978:Q4 and then form forecasts for horizons (quarters) $h = 1, \ldots, 16$. We then move forward one quarter and estimate the model from 1968:Q4 to 1979:Q1, and form forecasts for horizons 1-16. We continue in a similar fashion through the rest of the sample.[23]

Table 5 presents the ratios of the Root Mean Squared Error (RMSE) of the models HSM, M_1, M_2, and M_3 relative to the RMSE of the UC-SV model for 1, 2, 3, 4, 8, 12 and 16-quarters forecast horizons. In order to assess the statistical significance of the differences in the accuracy of the point forecasts in finite sample, we report the results of the Diebold-Mariano test for equal MSE, taking the UC-SV model as the benchmark. Following the recommendation of Clark and McCracken (2013), the Diebold-Mariano t-tests are computed using the adjusted variance developed in Harvey et al. (1997).[24] In order to evaluate the statistical significance, the tests are compared with the asymptotic critical values from the standard normal distribution.[25]

It is well-known that relatively simple statistical model or non-structural models such as the UC-SV model win forecast competitions against the different variants of structural Phillips curve models. However, as it can be seen from Table 5, when compared to the

[22]The point forecasts of the UC-SV model are computed as the median of the posterior distribution.

[23]Details on the construction of iterated and direct forecasts of the HSM model are reported in Appendix G, which also reports starting values and prior specification of the UC-SV model.

[24]We have also considered the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992) and obtained similar results.

[25]Since the UC-SV model is the benchmark model, the HSM baseline, M_1, M_2, M_3 are non nested under the null of equal forecast accuracy in finite sample.

UC-SV model, the null of equal accuracy is not rejected for the HSM and M_1 models for any of the forecast horizons considered. On the contrary, the null of equal accuracy in finite sample is rejected for 1 and 2-quarters ahead forecasts for the M_2 model respectively at the 1% and 5%, while for the M_3 model the null of equal accuracy is rejected at 1% at 8, 12 and 16-quarters ahead forecasts.[26]

Table 5: Ratios of RMSE (iterated forecasts, recursive scheme) to UC-SV model

Model	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 8$	$h = 12$	$h = 16$
HSM	1.029	1.009	1.023	1.020	1.018	0.986	0.968
M_1	1.056	1.119	1.132	1.137	1.081	0.889	0.750
M_2	1.943***	1.740**	1.516	1.318	0.909	0.612	0.529
M_3	1.020	1.045	1.037	1.088	1.162***	1.229***	1.252***

Notes: For HSM matrix $\hat{A}_{(t)}$ is estimated using the baseline VAR $Y_t = [lsi_t, y_t]$ with 4 lags; We use the Diebold-Mariano test for equal MSE applied to the forecast of each model relative to the benchmark UC-SV. Differences in accuracy that are statistically significant are denoted by *, **, ***, referring respectively to significance at 10%, 5%, 1%. The Diebold-Mariano test statistics are computed using the adjusted variance developed in Harvey et al. (1997).

The Diebold-Mariano statistic tests for equal forecast accuracy on average over the whole sample. In contrast, the Giacomini and Rossi (2010) fluctuations test focuses on the forecast accuracy at each point in time. In Appendix G we present the results of the Giacomini and Rossi (2010) fluctuations test (see Fig. G.2) for equal accuracy at each point in time of the HSM, M_1, M_2 and M_3 models relative to the benchmark UC-SV model. The test reveals some instabilities in the relative forecast accuracy of all the models which makes it difficult to conclude that one model is the best. In particular, Fig. G.2 (panel (a)) indicates that the HSM baseline model has more accurate forecasts than the UC-SV model for most of the 1980s, less accurate in the period following the 2007/8 financial crisis, and equal accuracy during the 1990s, 2000s and the recent years. Finally, in Appendix G we present rolling RMSEs (see Fig. G.1) using a weighted centered 15-quarter window (as

[26] The conclusions are similar when we use iterated point forecasts with a rolling window scheme (for the UC-SV, M_1, M_2, M_3 models) and direct point forecasts, both with recursive and rolling window scheme (for M_1, M_2, M_3 models, see Tables G.1, G.2 and G.3 in Appendix G).

in Stock and Watson (2009)). In particular, Fig. G.1 (panel (a)) shows that the baseline forecasts outperform the UC-SV model forecasts over most of the 1980s. Over the 1990s and 2000s the baseline HSM and UC-SV models forecasts are similar, while in the aftermath of the 2007/8 financial crisis the UC-SV model had smaller RMSE than the baseline HSM.

6 Survey data

So far we have presented evidence for behavioral heterogeneity and endogenous switching mechanism using macroeconomic data. In this section we focus on survey data on inflation expectations. Although a thorough analysis of micro-level data is beyond the scope of the paper, we provide some evidence for the behavioral mechanism described in the previous sections using observed average inflation expectations. Our measure of observed inflation expectations is given by the one-quarter ahead average expectations for the GDP deflator obtained from the Survey of Professional Forecasters (SPF). Following Coibion and Gorodnichenko (2012), Milani (2011) and Del Negro and Eusepi (2011) among others, we focus on SPF data because predictions of professional forecasters are consistently available at a quarterly frequency and such predictions, unlike the Michigan Survey of households forecasts, refer to explicitly defined variables such as the GDP deflator, and are therefore directly comparable with the measure of expectations extracted from model (12).[27] SPF data are available from 1968:Q4 and we compute expected inflation, denoted as $E_t^{spf}\pi_{t+1}$ as the log difference of the expected GDP deflator for the next quarter and the current quarter, averaged across forecasters.

Before proceeding with the analysis, a discussion on measurement issues with inflation expectations is warranted. Typically forecasters in the field only have the latest vintage of data available, while the econometrician often uses the final data vintage. Moreover, SPF forecasters provide their forecasts in the middle of the quarter and therefore they only have partial information about the current quarter. Although the issue of data revisions

[27] Branch (2004) fits a model with similar dynamic selection among three predictors, namely naive, adaptive and VAR expectations and costs associated with the adoption of each predictor, to the Michigan survey data on inflation expectations. The author finds support for switching behavior between different rules depending on the relative mean squared errors of the predictors.

represents an important challenge, we do not pursue it in this paper.[28] Therefore, in order to facilitate comparison with the results obtained in the previous sections, following Canova and Gambetti (2010) and Del Negro and Eusepi (2011) among others, we use inflation expectations together with revised data for macroeconomic variables and obtain the benchmark results of this session under the assumption that observed expectations are formed using an information set which includes the current quarter (see also Clark and Davig (2011) for a discussion on real-time data issues and data revisions in VARs.).

As a first exercise, we compare average SPF inflation expectations, i.e., $E_t^{spf}\pi_{t+1}$, with the time series of average expected inflation implied by model (12) (as estimated in Table 1), i.e., $\bar{E}_t\pi_{t+1} = n_{f,t}E_t^f\pi_{t+1} + (1 - n_{f,t})E_t^n\pi_{t+1}$. Fig. 3 plots the two time series.

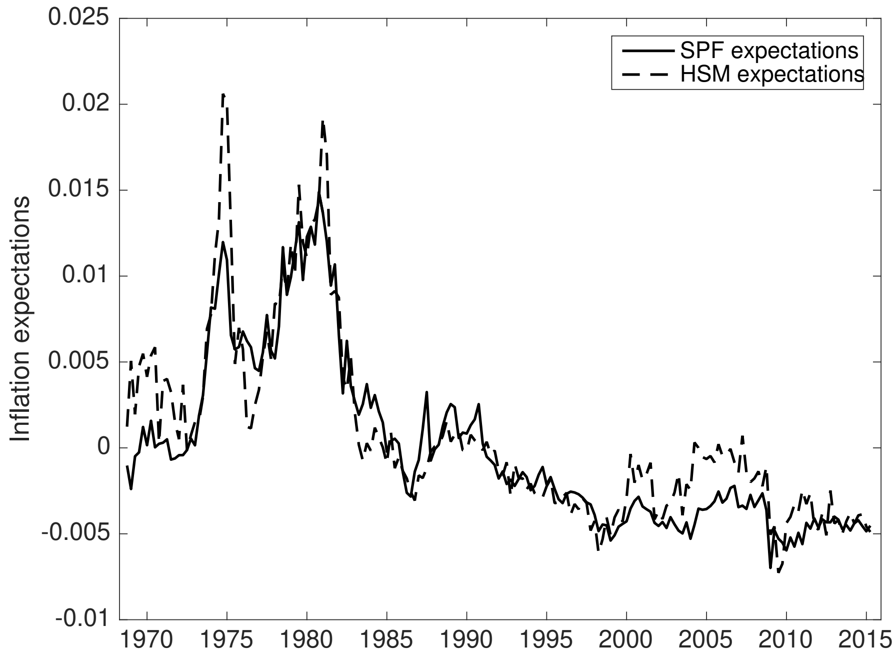

Figure 3: Inflation expectations: SPF data and HSM prediction

Fig. 3 shows that the model-implied expectations can reproduce, at least qualitatively,

[28]The "Real Time Data Set for Macroeconomists", made available by the Federal Reserve Bank of Philadelphia, reports vintages of the major macroeconomic data available at quarterly intervals in real time. However, the dataset does not include the labor share of income, while other variables are not available for the full sample considered in this paper.

the pattern of SPF expectations data. The correlation between the two series is 0.91.[29]

As a second exercise, we estimate the heuristic switching model directly on SPF average data and compare the results with the estimates in Table 1. In particular, we estimate the model

$$\begin{aligned}
E_t^{spf}\pi_{t+1} &= n_{f,t}E_t^f\pi_{t+1} + (1-n_{f,t})E_t^n\pi_{t+1} + \varepsilon_t \\
E_t^f\pi_{t+1} &= \gamma e_1'(I-\delta\hat{A}_{(t)})^{-1}\hat{A}_{(t)}Z_t \\
E_t^n\pi_{t+1} &= \pi_{t-1} \\
n_{f,t} &= \frac{1}{1+\exp\left(\beta\left(\frac{FE_{t-1}^f - FE_{t-1}^n}{FE_{t-1}^f + FE_{t-1}^n}\right)\right)} \\
FE_{t-1}^i &= \sum_{k=1}^{K}|E_{t-k-1}^i\pi_{t-k} - \pi_{t-k}|, \qquad \text{with } \ i = f, n,
\end{aligned}$$

using the baseline VAR specification. The estimation results are reported in Table 6.

Table 6: NLS estimates of HSM with SPF data

Parameter	β	γ
Estimate	1.995***	0.002***
Std. error	0.355	0.0002
R^2 from SPF Equation	0.83	

Notes: Standard errors are computed using White's heteroskedasticity-consistent covariance matrix estimator (HCCME). *, **, *** denote significance at the 10%, 5%, and 1% level.

The estimate of parameter γ using SPF data is positive and statistically significant, with a magnitude comparable to the estimate in Table 1. Although the estimate of parameter β is slightly smaller than the estimate obtained using aggregate data, suggesting slower switching dynamics in SPF data, we do observe a positive and significant β. This means that survey data support the hypothesis of dynamic switching that depends on past forecasting errors. The time series of $n_{f,t}$ extracted using both macroeconomic and SPF data are displayed in Fig. 4.

[29] For the sake of completeness, we also compare average SPF expectations data with expectations paths generated by models in which $n_{f,t} = 0, 1$. The R^2 (log-likelihood) for the HSM model is 0.83 (894) and for the models with $n_{f,t} = 0, 1$ they are respectively 0.79 (874) and 0.15 (742).

Finally, in order to provide some evidence that the time-variation in the weights extracted by model (12) is associated with changes in the behavior of actual forecasters, we perform the following exercise. First, we estimate a model with constant weights of fundamentalists and random walk believers using SPF data, i.e., $n_{f,t} = n_f$. Second, we test for unknown multiple structural breaks in the estimated weight $\hat{n}_f$ using the sequential Sup-Wald test (robust to autocorrelation and heteroskedasticity) proposed by Bai and Perron (1998). We first test for zero breaks against the alternative of one break. If the null is rejected, we then test for an additional break in the subsample before and after the first break, and so on until the alternative of an additional break is not rejected. The Bai and Perron procedure detected the following break dates: 1972:Q4, 1974:Q4, 1983:Q1, 1993:Q1, 1999:Q1, 2005:Q3, 2009:Q4, 2011:Q2. The break dates are all significant at 1%, except for the breaks in 2005:Q3 and 1993:Q1 which are significant at 5%. Third, we plot in Fig. 4 the estimated breaks against the time series of the weight $n_{f,t}$ extracted by model (12) from macroeconomic data. The results displayed in Fig. 4 show that indeed the

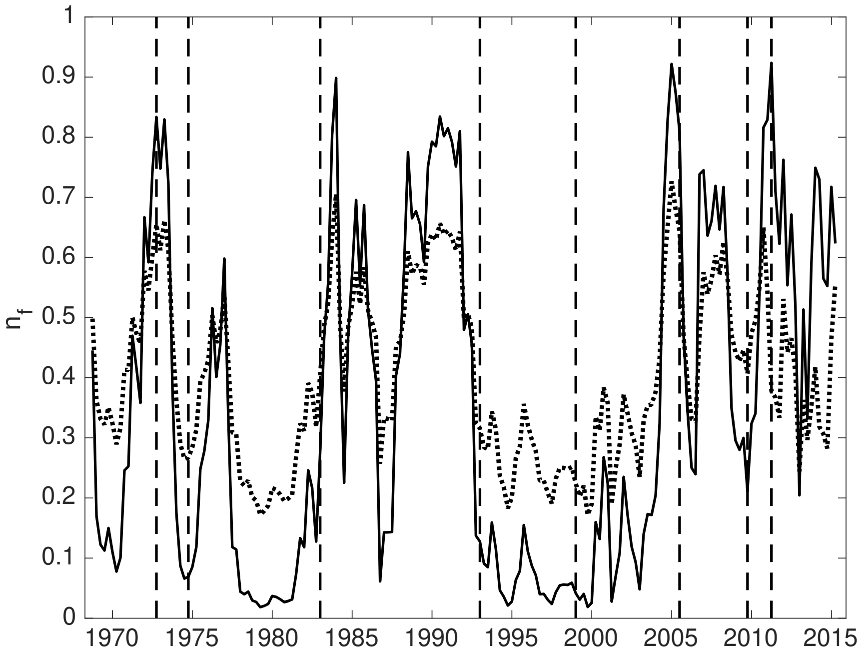

Figure 4: Estimated structural breaks on SPF data (vertical dashed lines) and HSM weights extracted from aggregate data (solid line) and SPF data (dashed line).

estimated structural changes in survey data on actual forecasting behavior correspond to large movements in the weights of forward- and backward-looking terms in the NKPC. For example, the estimated breaks in 1972:Q4 and 1974:Q4 correspond to a phase in which the weight of fundamentalists oscillates from about 0.1 to 0.8, and then back to about 0.05; or, as another example, the estimated breaks in 1983:Q1 and 1993:Q1 match respectively a change of the weight of forward-looking behavior from around 0.1 to 0.9 and from 0.8 to 0.1.

Overall, the results presented in this section support the idea of an heuristic switching mechanism based on past performance in survey data on actual inflation expectations. Moreover, the fluctuations in weights of forward- vs. backward-looking behavior extracted by the model using macroeconomic data seem to be matched by some changes in the predicting behavior of actual forecasters.

7 Conclusions

Over the past few decades it has become accepted that the purely forward-looking NKPC cannot account for the degree of inflation inertia observed in the data. In response, the profession has increasingly adopted hybrid models in which lagged inflation is allowed to have an explicit role in pricing behavior. This reformulation of the basic sticky-price model has recently provoked a heated debate as to the extent of forward- versus backward-looking behavior, with little consensus after years of investigation. Most of the empirical studies on the topic take the distribution of heterogeneous pricing behavior as fixed and exogenously given. Recent works on structural stability in short-run inflation dynamics in the U.S. have provided statistical evidence of multiple structural breaks in the relative weights of forward- and backward-looking firms. Moreover, empirical studies based on survey data as well as experimental data, provided evidence that the proportions of heterogeneous forecasters evolve over time as a reaction to past forecast errors. In the light of this empirical evidence, we have proposed a model of monopolistic price setting with nominal rigidities and endogenous evolutionary switching between different forecasting strategies according to their relative past performances. Importantly, in light of the recent criticisms to model-consistent RE in the NKPC on both theoretical and empirical grounds, heterogeneous firms in our model hold optimizing behavior given their *subjective* expectations of future inflation. In our stylized framework, fundamentalist firms believe in a present-value relationship between inflation and real marginal costs, as predicted by standard RE models, while naive firms use a simple rule of thumb to forecast future inflation. Although with a different behavioral interpretation, our measure of fundamental expectation is similar to the measure of forward-looking expectations in commonly estimated RE models, while the expectations of random walk believers account for the lagged value of inflation in the hybrid specification of the NKPC. The difference with traditional tests of sticky-price models arises from the introduction of time-varying weights and endogenous switching dynamics.

We estimated our behavioral model of inflation dynamics on quarterly U.S. inflation data from 1968:Q4 to 2015:Q2. Our estimation results show statistically significant behavioral heterogeneity and substantial time variation in the weights of forward- and backward-looking price setters. The data gave considerable support for the parameter restrictions implied by our theory. In particular, the intensity of choice was found to be positive, indicating that agents switch towards the better performing rule according to its past performance, and inflation was positively affected by real marginal costs. Moreover, the analysis of survey data on inflation expectations provides additional evidence for the behavioral switching mechanism.

Our results suggest that large shocks, such as the oil crises, pushed inflation away from the fundamental, triggering therefore a large share of agents to adopt random walk beliefs. Moreover, during the Great Moderation, forecasts obtained in accordance to a Phillips curve relationship between inflation and real activity, did not perform better than simple univariate models, such as the naive rule.

Our findings have important monetary policy implications. Recent papers have shown that multiple equilibria and complex dynamics can arise in New Keynesian models under dynamic predictor selection, even if the model under RE has a unique stationary solution. Given the statistical evidence found in our empirical results for heterogeneous expectations and evolutionary switching, determinacy under RE may not be a robust recommendation and monetary policy should be designed to account for potentially destabilizing heterogeneous expectations.

References

Adam, K. (2007). Experimental Evidence on the Persistence of Output and Inflation. *The Economic Journal 117*(520), 603–636.

Allen, H. and M. Taylor (1990). Charts, noise and fundamentals in the london foreign exchange market. *The Economic Journal 100*, 49–59.

Andrews, D. W. K. and J. C. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica 60*, 953 – 966.

Anufriev, M., T. Assenza, C. Hommes, and D. Massaro (2013). Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics 17*, 1574–1604.

Arifovic, J., J. Bullard, and O. Kostyshyna (2012). Social learning and monetary policy rules. *Economic Journal 123*, 38 – 76.

Assenza, T., P. Heemeijer, C. Hommes, and D. Massaro (2011). Individual expectations and aggregate macro behavior. CeNDEF Working Paper 2011-1, University of Amsterdam.

Atkeson, A. and L. E. Ohanian (2001). Are Phillips curves useful for forecasting inflation? *Quarterly Review*, 2–11.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica 66*(1), 47–78.

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

Branch, W. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations. *The Economic Journal 114*(497), 592–621.

Branch, W. and B. McGough (2009). Monetary policy in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control 33*, 1036–1051.

Branch, W. and B. McGough (2010). Dynamic predictors selection in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control 34*(8), 1492–1508.

Branch, W. A. and G. W. Evans (2016). Unstable inflation targets. *Journal of Money, Credit and Banking*. Forthcoming.

Brock, W. A. and C. H. Hommes (1997). A rational route to randomness. *Econometrica 65*(5), 1059–1095.

Campbell, J. and R. Shiller (1987). Cointegration and tests of present value models. *Journal of Political Economy 95*, 1062–1088.

Canova, F. and L. Gambetti (2010). Do Expectations Matter? The Great Moderation Revisited. *American Economic Journal: Macroeconomics 2*(3), 183–205.

Carriero, A. (2008). A simple test of the new keynesian phillips curve. *Economics Letters 100*, 241–244.

Carroll, C. (2003). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics 118*, 269–298.

Castle, J., J. Doornick, D. Hendry, and R. Nymoen (2014). Mis-specification testing: Non-invariance of expectations models of inflation. *Econometric Reviews 33*, 553–574.

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

Clark, T. E. and T. Davig (2011). Decomposing the declining volatility of long-term inflation expectations. *Journal of Economic Dynamics and Control 35*(7), 981 – 999.

Clark, T. E. and M. W. McCracken (2013). Advances in forecast evaluation, volume 2. In G. Elliott and A. Timmermann (Eds.), *Hanbook of economic forecasting*. Elsevier.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Davidson, R. and J. G. MacKinnon (1981). Several tests for model specification in the presence of alternative hypothesis. *Econometrica 49*, 781–793.

Davidson, R. and J. G. MacKinnon (2009). *Econometric Theory and Methods*. Oxford University Press, Oxford.

De Grauwe, P. (2011). Animal spirits and monetary policy. *Economic Theory 47*, 423–457.

Del Negro, M. and S. Eusepi (2011). Fitting observed inflation expectations. *Journal of Economic Dynamics and Control 35*(12), 2105–2131.

Diks, C. and R. van der Weide (2005). Herding, a-synchronous updating and heterogeneity in memory in a cbs. *Journal of Economic Dynamics and Control 29*, 741–763.

Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Frankel, J. and K. Froot (1990). Chartists, fundamentalists and trading in the foreign exchange market. *American Economic Review 80*(2), 181–185.

Frankel, J. and K. Froot (1991). Chartists, fundamentalists and the demand for dollars. In A. Courakis and M. Taylor (Eds.), *Private Behaviour and Government Policy in Interdependent Economies*. Oxford: Clarendon.

Fuhrer, J. (1997). The (un)importance of forward-looking behavior in price specification. *Journal of Money, Credit and Banking 29*, 338–350.

Fuhrer, J. and G. Moore (1995). Inflation persistence. *Quarterly Journal of Economics 110*, 127–59.

Galí, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics 44*, 195–222.

Giacomini, R. and B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics 25*, 595 – 620.

Groen, J. J. J., R. Paap, and F. Ravazzolo (2013). Real-Time Inflation Forecasting in a Changing World. *Journal of Business & Economic Statistics 31*(1), 29–44.

Hachem, K. and J. Wu (2014). Inflation announcement and social dynamics. Chicago Booth Working Paper 13-76.

Hall, A., S. Han, and O. Boldea (2012). Inference regarding multiple structural changes in linear models with endogenous regressors. *Journal of Econometrics 170*, 281–302.

Harvey, D. I., S. J. Leybourne, and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting 13*, 281–291.

Hendry, D. and G. Mizon (2010). On the mathematical basis of inter-temporal optimization. Economics Series Working Papers 497, University of Oxford.

Hendry, D. and G. Mizon (2014). Unpredictability in economic analysis, econometric modelling and forecasting. *Journal of Econometrics 182*, 186–195.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies 18*(3), 955–980.

Ito, T. (1990). Foreign exchange rate expectations: Micro survey data. *American Economic Review 80*(3), 434–449.

Kim, C. and Y. Kim (2008). Is the backward-looking component important in a new keynesian phillips curve? *Studies in Nonlinear Dynamics and Econometrics 12*.

Kleibergen, F. and S. Mavroeidis (2009). Weak instrument robust tests in gmm and the new keynesian phillips curve. *Journal of Business & Economic Statistics 27*(3), 293–311.

Kreps, D. (1998). Anticipated utility and dynamic choice. In *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983-1997*. Cambridge University Press.

Kurmann, A. (2007). Var-based estimation of euler equations with an application of new keynesian pricing. *Journal of Economic Dynamics and Control 31*, 767–796.

Kurz, M., G. Piccillo, and H. Wu (2013). Modeling diverse expectations in an aggregated new keynesian model. *Journal of Economic Dynamics and Control 37*, 1403–1433.

Lindé, J. (2005). Estimating new-keynesian phillips curves: A full information maximum likelihood approach. *Journal of Monetary Economics 52*, 1135–1149.

Madeira, C. and B. Zafar (2015). Heterogeneous inflation expectations and learning. *Journal of Money, Credit and Banking 47*, 867–896.

Madeira, J. (2014). Overtime labor, employment frictions, and the new keynesian phillips curve. *The Review of Economics and Statistics 96*(3), 767–778.

Madeira, J. (2015). Firm-specific capital, inflation persistence and the sources of business cycles. *European Economic Review 74(C)*, 229–243.

Mankiw, N., R. Reis, and J. Wolfers (2003). Disagreement about inflation expectations. In M. Gertler and K. Rogoff (Eds.), *NBER Macroeconomics Annual*. National Bureau of Economic Research.

Manski, C. and D. McFadden (1981). Alternative estimators and sample designs for discrete choice analysis. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data*. MIT Press.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–298.

Milani, F. (2011). Expectation shocks and learning as drivers of the business cycle*. *The Economic Journal 121*(552), 379–401.

Pfajfar, D. and E. Santoro (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior and Organization 75*(3), 426–444.

Pfajfar, D. and B. Zakelj (2014). Experimental evidence on inflation expectation formation. *Journal of Economic Dynamics and Control 44*, 147–168.

Romer, D. (2011). *Advanced Macroeconomics*. McGraw-Hill.

Rudd, J. and K. Whelan (2005a). Does the labor's share drive inflation? *Journal of Money, Credit and Banking 37*, 297–312.

Rudd, J. and K. Whelan (2005b). New tests of the new keynesian phillips curve. *Journal of Monetary Economics 52*, 1167–1181.

Rudd, J. and K. Whelan (2006). Can rational expectations sticky price models explain inflation dynamics? *American Economic Review 96*, 303–320.

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

Sbordone, A. (2002). Prices and unit labor costs: A new test of price stickiness. *Journal of Monetary Economics 49*(2), 265–292.

Sbordone, A. (2005). Do expected future marginal costs drive inflation dynamics? *Journal of Monetary Economics 52*, 1183–1197.

Smets, F. and R. Wouters (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review 97*(3), 586–606.

Stock, J. H. and M. W. Watson (2007). Why has u.s. inflation become harder to forecast? *Journal of Money, Credit and Banking 39*, 3–33.

Stock, J. H. and M. W. Watson (2009). Phillips curve inflation forecasts. In J. Fuhrer, K. Y., L. J., and O. G. (Eds.), *Understanding Inflation and the Implication for the Monetary Policy*. Cambridge MIT press.

Tuinstra, J. and F. Wagener (2007). On learning equilibria. *Economic Theory 30*(3), 493–513.

Woodford, M. (2001). The taylor rule and optimal monetary policy. *American Economic Review 91*, 232–237.

Zhang, C., D. Osborn, and D. Kim (2008). The new keynesian phillips curve: From sticky inflation to sticky prices. *Journal of Money, Credit and Banking 40*, 667–699.