Macroeconomic Dynamics, **20**, 2016, 1073–1100. Printed in the United States of America.
doi:10.1017/S1365100514000741

RECURSIVE INATTENTIVENESS WITH HETEROGENEOUS EXPECTATIONS

Lena Dräger
University of Hamburg

We model agents' endogenous updating of information sets over time under changing macroeconomic conditions. Building on sticky information models, the degree of inattentiveness is endogenized by allowing agents to choose between a costly full-information predictor and a costless sticky-information predictor. This is modeled as a choice between discrete alternatives under rational inattention. Recursive simulation shows that the dynamic equilibrium paths of aggregate variables are highly persistent and match the moments of U.S. data better than a model with fixed sticky information or with sticky prices, especially with regard to higher moments and the degree of persistence. Predictors are chosen in line with the predictions from rational inattention models, as the aggregate degree of attentiveness increases with rising variance of the forecast variable. Moreover, the model can generate hump-shaped impulse responses of inflation to a monetary policy shock if the degree of inattentiveness is sufficiently high.

Keywords: Sticky Information, Recursive Inattentiveness, Rational Inattention, Heterogeneous Expectations, DSGE Models

1. INTRODUCTION

How do agents form expectations regarding future inflation or output if they are subject to information constraints? Sticky information models of Mankiw and Reis (2002, 2003, 2007), as well as Ball et al. (2005), propose that agents have a given probability $\overline{\lambda}$ of being able to update their information set each period, which is motivated by costs related to the acquisition and processing of information. Information thus becomes "sticky" and although agents are assumed to be rational, they may nevertheless be forced to form expectations based on

An earlier working paper version was circulated under the title "Why Don't People Pay Attention? Endogenous Sticky Information in a DSGE Model." The author would like to thank the Associate Editor, two anonymous referees, Klaus Adam, Ulrich Fritsche, Ingrid Größl, Tom Holden, Jan-Oliver Menz, Alexander Meyer-Gohde, Ricardo Reis, Jan-Egbert Sturm, and Mirko Wiederholt, as well as conference and seminar participants at the Royal Economic Society 2011 conference, the European Economic Association 2011 conference, the German Economic Association 2012 conference, the German Bundesbank, the New School for Social Research, New York, ZEW Mannheim, and the University of Hamburg for very helpful comments and Paul De Grauwe for sharing his code. I thankfully acknowledge financial support from the German Research Foundation (Grant FR 2677/2-2). All remaining errors are mine. Address correspondence to: Lena Dräger, Department of Economics, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany; e-mail: Lena.Draeger@wiso.uni-hamburg.de.

© 2015 Cambridge University Press 1365-1005/15

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

outdated information. As in the Calvo (1983) pricing model, the probability $\overline{\lambda}$ of updating information is given exogenously and, thus, not related to a specification of the acquisition and processing costs.[1]

In a related class of models, the concept of "rational inattention" is proposed by Sims (2003) as a more general approach to modeling agents' limitations in the processing of information. Similarly to models with sticky information, agents are assumed to be rational, but nevertheless have a limited capacity for processing information. Thus, agents rationally choose to remain inattentive to certain bits of information.

In addition to these theoretical advances, there exists ample evidence in the literature that expectations from survey as well as experimental data show persistent heterogeneity and frequent switching between different predictors [e.g., Thomas (1999); Hommes et al. (2005); Hommes (2011); Pfajfar and Zakelj (2011)] and that the degree of attentiveness varies significantly over time. Branch (2007) compares the standard sticky information model with a version where agents may choose their recursive updating frequencies, derived in Branch (2004). The author finds that a model with time-varying updating frequencies provides a better fit to survey data on consumers' inflation expectations from the Michigan Survey of Consumers, where the proportions of agents choosing each frequency vary over time. Cornea et al. (2012) estimate a model with time-varying switching between fundamentalist and naive predictors for inflation dynamics in the United States and report that the data imply substantial time variation in beliefs. Using the rotating panel dimension of the survey, where some households are reinterviewed after six months, we can moreover directly calculate the share of households that updated their qualitative inflation expectations between interviews, i.e., the average degree of attentiveness regarding inflation [see Dräger and Lamla (2012)]. The time-varying share of households with updated short-run qualitative inflation expectations from the Michigan survey is shown in Figure 1, where the shaded areas denote recession periods as defined by the NBER. The updating share varies cyclically between about 20% and 60% and rises during recessions, implying that consumers increase attention to inflation during these periods.

In line with the empirical evidence on time-varying attentiveness, the question analyzed in this paper therefore is as follows: How are agents' costly information updates related to changes in the macroeconomic environment? We propose a link between the sticky information and the rational inattention model and endogenize the probability of obtaining the new information set, $\overline{\lambda}$, as the outcome to a discrete choice under rational inattention between a costly full information predictor and a costless sticky information predictor. The main novelty of the paper is that our approach yields a time-varying and state-dependent probability λ_t of choosing forecasts with full information in a general equilibrium setting. This allows us to address the empirical evidence of time-varying inattentiveness and to analyze feedback effects between the aggregate degree of inattentiveness and the evolution of the model economy over time. Note that by allowing that all agents may choose

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

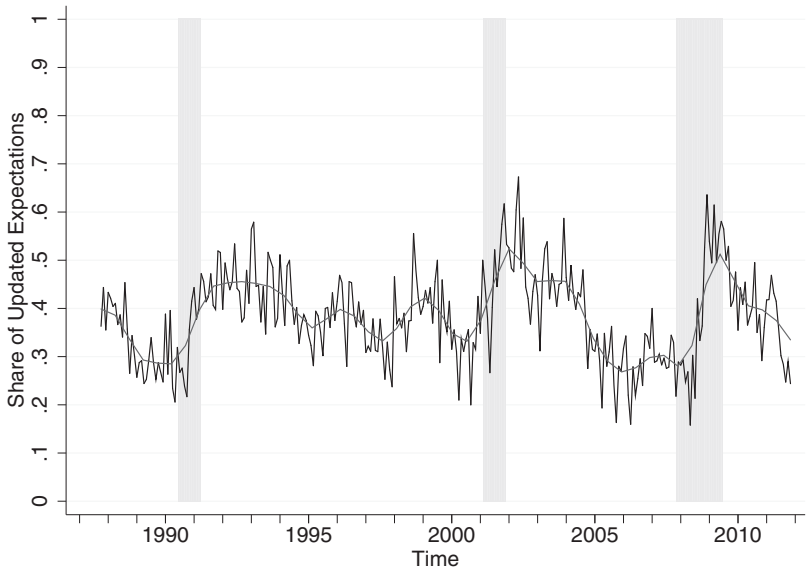


FIGURE 1. The degree of attentiveness regarding inflation from U.S. households microdata. *Source:* Microdata from the University of Michigan Survey of Consumers, own calculations. The degree of attentiveness is calculated as the share of households that change their inflation expectations within six months. Shaded areas denote recession periods as defined by the NBER.

the costly full information predictor, our model nests the full information rational model as a limiting case.

In our model with recursive inattentiveness, agents face two optimization problems: On one hand, households and firms optimize expected utility and profits, subject to their individual expectations. On the other hand, agents in each period choose the optimal predictor under rational inattention, subject to a discrete choice between a costly full information predictor and a costless sticky information predictor. This is modeled building on Matějka and McKay (2011), who show that, under the assumption of a priori symmetric options, the multinomial logit model emerges as the optimal strategy of a decision maker with discrete choices under rational inattention. We apply their approach to the choice between predictors and then solve the model recursively each period, yielding a dynamic equilibrium path for aggregate output, inflation, and nominal interest rates with a time-varying probability of choosing the full information predictor λ_t.

We compare numerical simulations of aggregate variables from our model with results from a model with sticky information and a fixed degree of inattentiveness $\overline{\lambda}$, and from a model with sticky prices and full information rational expectations. As a benchmark, we use detrended U.S. data on real output growth, CPI inflation,

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

and short-run nominal interest rates. Generally, our model with recursive inattentiveness matches the moments of U.S. data best, especially with respect to higher moments such as skewness and kurtosis. Moreover, the high degree of persistence generally found in the U.S. data is matched best by our model. Especially in the case of inflation, our model yields strong endogenous persistence, whereas neither the sticky information nor the sticky price model can produce significant inflation persistence, because we omit autocorrelated shocks and a lagged inflation term. In that sense, our model also relates to the models with boundedly rational inflation expectations in Ball (2000), Lansing (2009), and De Grauwe (2011), who find that their models yield strong persistence in inflation. However, our model has the advantage of nesting full information expectations as a special case.

Analyzing agents' degree of attentiveness to output and inflation, we find that attentiveness is positively correlated with the variance of the variable to be forecasted. Hence, agents behave in line with the predictions of rational inattention models, where a higher variance of the forecasted variable is a signal to devote more attention to that variable [Mackowiak and Wiederholt (2009)]. Although the recursively simulated degree of attentiveness to inflation is more volatile, it matches the amplitude and cyclical behavior of the empirical measure from households' inflation expectations, shown in Figure 1, reasonably well.

Finally, regarding the effects of a monetary policy shock on aggregate inflation and output, we find that the model can replicate the stylized fact of a hump-shaped impulse response of inflation if the degree of inattentiveness is sufficiently high, whereas impulse responses of output are persistent, but not hump-shaped. However, in periods with a high degree of attentiveness to output and inflation, impulse responses are stronger and peaked.

Our model with recursive inattentiveness is related to the literature on heterogeneous expectations, which frequently employs the multinomial logit model for probabilistic choices between different predictors. Starting with the seminal paper by Brock and Hommes (1997), the multinomial logit is usually motivated by its derivation from a random utility model and is also frequently applied in discrete choice theory [see for instance Manski and McFadden (1981)]. The models of Brazier et al. (2008) and De Grauwe (2010) employ the multinomial logit to analyze switching over time between optimistic/pessimistic heuristics or between predictors based on lagged inflation and the official inflation target, respectively. Recently, Anufriev et al. (2013) used a similar model with heterogeneous expectations to study the role of monetary policy in stabilizing the economy.

In addition to the models with heterogeneous expectations, our approach is related to a number of models in the literature on sticky information and on rational inattention. In a sticky information model, as in Ball et al. (2005), Branch et al. (2006, 2009) derive the optimal degree of inattentiveness by firms. In their model, firms choose their optimal information updating probability in relation to an economywide degree of inattentiveness and subject to a fixed information cost.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

The authors show that a symmetric Nash equilibrium degree of inattentiveness exists. Our approach differs from theirs in that we analyze changes in endogenous inattentiveness over time and via the recursive solution we allow for feedback from the time-varying degree of inattentiveness to the model evolution.

From the literature on rational inattention, the model in Woodford (2009) is most closely related to our approach. Woodford (2009) presents a model of state-dependent pricing by firms with rational inattention, where Bayesian firms face costs related both to the updating of information and to receiving a signal on possible normalized prices. The probability of updating information is governed by a hazard function, which is given as a function of the firm's normalized price. The model in Woodford (2009) differs from ours, because it focuses on the problem of a single firm, whereas in this paper we study a general equilibrium economy to analyze the interdependence between changes in the degree of inattentiveness and the evolution of the macroeconomy. Other relevant approaches include those of Adam (2007, 2009), Mackowiak and Wiederholt (2009, 2010), and Paciello and Wiederholt (2014), as well as Hahn (2014), who present models with rational inattention of firms, solving for the equilibrium degree of inattention and analyzing optimal monetary policy and the role of policy transparency.

The remainder of the paper is structured as follows. The model is derived and discussed in Section 2. Results of the model simulations with recursive inattentiveness are then given in Section 3. Finally, Section 4 summarizes and concludes.

2. THE MODEL

Our model takes the form of a New Keynesian DSGE model as in Clarida et al. (1999) or Woodford (2003), but includes flexible prices and heterogeneous expectations. We assume flexible prices in line with the sticky information models in Mankiw and Reis (2002, 2003, 2007), where the only rigidity comes from stickiness in information. The model economy is populated by a large number of infinitely lived households with identical preferences and endowments, which may differ with respect to their expectations regarding inflation and output. Households maximize expected utility subject to their budget constraint and to heterogeneous expectations. A large number of firms produce differentiated consumption goods under monopolistic competition and flexible prices. In addition, we assume that firms are owned by households and maximize expected profits subject to their production function, heterogeneous expectations, and households' demand.

The optimization problem regarding the optimal predictor is assumed to take place after the maximization of expected utility or profits. Thus, heterogeneous expectations are introduced ex post into aggregate macroeconomic relationships, after the derivation of the aggregate New Keynesian IS curve and the sticky information Phillips curve. This approach is quite common in the literature on heterogeneous expectations, and also in the learning literature; see for instance Branch and Evans (2011), De Grauwe (2011), and Branch et al. (2013), as well as Berardi and Duffy (in press).[2]

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

2.1. Heterogeneous Expectations

In the sticky information literature, the aggregate forecast $\widehat{E}_t$ consists of an infinite sum of expectations formed under current and past information sets, weighted with the fixed exogenous probability of obtaining new information $\overline{\lambda}$ [Mankiw and Reis (2002)]:

$$\begin{aligned}\widehat{E}_t(x) &= \overline{\lambda}\sum_{j=0}^{\infty}\left(1-\overline{\lambda}\right)^j E_{t-j}(x)\\ &= \overline{\lambda}E_t(x) + \left(1-\overline{\lambda}\right)\overline{\lambda}E_{t-1}(x) + \left(1-\overline{\lambda}\right)^2\overline{\lambda}E_{t-2}(x)\ldots. \qquad (\mathbf{1})\end{aligned}$$

Allowing for a time-varying probability λ_t of obtaining the full information forecast, i.e., a time-varying degree of attentiveness, the aggregate forecast is derived as a convex combination of the full information forecast and past expectations under sticky information:

$$\begin{aligned}\widehat{E}_t(x) &= \lambda_t E_t(x) + (1-\lambda_t)\widehat{E}_{t-1}(x)\\ &= \lambda_t E_t(x) + (1-\lambda_t)\lambda_{t-1}E_{t-1}(x)\\ &\quad + (1-\lambda_t)(1-\lambda_{t-1})\lambda_{t-2}E_{t-2}(x) + \ldots. \qquad (\mathbf{2})\end{aligned}$$

Because the expression in (2) requires following the dynamics of λ_t up to $(t-\infty)$, past expectations under sticky information, $\widehat{E}_{t-1}(x)$, are approximated in our model by setting the weights $\lambda_{t-j} = \overline{\lambda} \quad \forall j = 1, 2, \ldots$. The weight $\overline{\lambda}$ may also be interpreted as the past average degree of attentiveness.[3] Thus, we assume that past changes in the probability of obtaining the full information forecast are not reflected in $\widehat{E}_{t-1}$. Note, however, that this simplifying assumption affects only the weights on past expectations, whereas the forecasts $E_{t-j}(x)$ are recursively updated each period and thus account for the effect of changes in λ_t on the realized forecasts up to $t-j$.

Agents in the model thus face a discrete choice between two predictors: The full information predictor denoted as $E_t^{\mathrm{FI}} = E_t$, on one hand, and the sticky information predictor given by $E_t^{\mathrm{SI}} = \overline{\lambda}\sum_{j=0}^{\infty}(1-\overline{\lambda})^j E_{t-1-j}$, on the other hand. Aggregate forecasts are then obtained as the heterogeneous expectations index $\widetilde{E}_t(x)$:

$$\begin{aligned}\widetilde{E}_t(x) &= \lambda_t E_t^{\mathrm{FI}}(x) + (1-\lambda_t)\, E_t^{\mathrm{SI}}(x)\\ &= \lambda_t E_t(x) + (1-\lambda_t)\,\overline{\lambda}\sum_{j=0}^{\infty}\left(1-\overline{\lambda}\right)^j E_{t-1-j}(x)\\ &= \lambda_t E_t(x) + (1-\lambda_t)\,\overline{\lambda}E_{t-1}(x) + (1-\lambda_t)\left(1-\overline{\lambda}\right)\overline{\lambda}E_{t-2}(x) + \ldots. \qquad (\mathbf{3})\end{aligned}$$

As argued in Branch (2007), switching between predictors under full or outdated information is logically consistent if information is costly and if agents obtain

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

the forecast, but not the method. Information costs related to the obtaining and processing of new information provide the general motivation for sticky information [see Ball et al. (2005); Reis (2006a, 2006b); and Branch et al. (2009)]. In our model, information costs may be thought of as the cost of acquiring and incorporating the most recent expert forecast. The sticky information forecast may thus be regarded as being obtained from nonexperts and includes own past forecasts subject to some degree of memory loss.[4]

If agents switch from using the sticky information forecast to using the full information forecast at the beginning of period t, they pay the information cost and obtain the expert forecast incorporating all information up to period t. Because as a result of the recursive structure of the model forecasts are reestimated each period, when switching from sticky to full information it is optimal to use only the most recent, full information, forecast. However, if agents switch from using the full information to the sticky information forecast at the beginning of period t, they are assumed to suffer from a certain memory loss because their own past forecast in period $t-1$ enters only with weight $\bar{\lambda}$ into the sticky information forecast. Following Branch (2007), we argue that this is nevertheless consistent if agents obtain forecasts from external sources, because then they do not discard important information when forming expectations themselves, but rather suffer from memory loss regarding their obtained forecasts.[5]

In order to be able to derive the aggregate heterogeneous expectations index $\widetilde{E}_t(x)$ in equation (3), we furthermore make the simplifying assumptions that, on the average, agents do not expect changes in the composition of average future expectations $\widetilde{E}_{t+j}(x)$ and that agents have the same expectations regarding limiting real income. This means that the law of iterated expectations is assumed to hold also across predictors, thus ruling out an effect of higher-order beliefs [Branch and McGough (2009)].[6]

2.2. Aggregate Supply and Demand

Aggregate demand in our model is derived from the standard utility maximization problem in DSGE models, where we allow for heterogeneous expectations. Because the demand side of the model follows a standard DSGE setup, we omit detailed derivations and refer the reader to the survey in Clarida et al. (1999) for the derivation of the New Keynesian IS curve. From households' utility maximization problem, we get the individual log-linearized Euler equation

$$\hat{c}_{t,j} = E_t^j \hat{c}_{t+1,j} - \frac{1}{\sigma}(\hat{i}_t - E_t^j \pi_{t+1}), \tag{4}$$

where $\hat{c}_{t,j}$ denotes the consumption of agent j, $\hat{i}_t$ is the nominal interest rate, π_{t+1} is the next period's inflation rate (all variables with a circumflex are measured in deviations from the steady state), and the individual expectations operator is E_t^j. Note that the specification of the Euler equation in (4) implicitly assumes that all agents can observe deviations of the current nominal interest rate from its steady

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

state in $\hat{i}_t$. This assumption can be interpreted as representing the willingness of all agents to pay for the costs related to processing information on changes in the interest rate set by the central bank.

The aggregation of individual Euler equations in (4) thus calls for an aggregation of heterogeneous predictors in E_t^j. We apply the aggregate heterogeneous expectations index $\widetilde{E}_t$, as defined in equation (3): $\widetilde{E}_t(x) = \lambda_t E_t^{\mathrm{FI}}(x) + (1-\lambda_t)\, E_t^{\mathrm{SI}}(x)$. Noting that market clearing implies $\hat{y}_t = \hat{c}_t + u_t$,[7] we thus get an expression for aggregate output $\hat{y}_t$ in terms of the New Keynesian IS curve with heterogeneous expectations:

$$\begin{aligned}\hat{y}_t &= \widetilde{E}_t \hat{y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \widetilde{E}_t \pi_{t+1}) + u_t \\ &= \lambda_t E_t\left[\hat{y}_{t+1} + \frac{1}{\sigma}\pi_{t+1}\right] \\ &\quad + (1-\lambda_t)\overline{\lambda}\sum_{j=0}^{\infty}(1-\overline{\lambda})^j E_{t-1-j}\left[\hat{y}_{t+1} + \frac{1}{\sigma}\pi_{t+1}\right] - \frac{1}{\sigma}\hat{i}_t + u_t. \qquad (5)\end{aligned}$$

Aggregate supply is derived as in Mankiw and Reis (2002) and Ball et al. (2005), assuming flexible prices and allowing for sticky information and heterogeneous expectations. Again, we refer to Mankiw and Reis (2002) and Ball et al. (2005) for detailed derivations. From firms' profit maximization problem, we get firm i's optimal price under flexible prices $\hat{p}_{t,i}^*$, given their individual expectations E_t^i:

$$\hat{p}_{t,i}^* = E_t^i\left[\hat{p}_t + \hat{\varphi}_t\right], \qquad (6)$$

where $\hat{p}_t$ denotes the aggregate price index and $\hat{\varphi}_t$ measures real marginal costs. Assuming that all agents are subject to sticky information, Ball et al. (2005) derive from (6) the sticky information Phillips curve with the output gap instead of real marginal costs:

$$\begin{aligned}\pi_t &= \frac{\psi\overline{\lambda}}{1-\overline{\lambda}}\left(\hat{y}_t - \hat{y}_t^n\right) + \overline{\lambda}\sum_{j=0}^{\infty}(1-\overline{\lambda})^j E_{t-1-j}\left[\pi_t + \psi\Delta\left(\hat{y}_t - \hat{y}_t^n\right) + \Delta e_t\right] \\ &\quad + \frac{\overline{\lambda}}{1-\overline{\lambda}}e_t, \qquad (7)\end{aligned}$$

where the output gap $(\hat{y}_t - \hat{y}_t^n)$ is defined as the deviation of output from the full information, flexible price output $\hat{y}_t^n$, which is driven by an i.i.d. technology shock. Note that e_t is an i.d.d. cost-push shock resulting, for instance, from unexpected wage or tax changes and $\overline{\lambda}$ is the exogenous probability of an information update. Because our aim is to endogenize $\overline{\lambda}$, heterogeneous expectations are introduced into the aggregate supply equation in (7) as before, so that we get aggregate

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

heterogeneous expectations formed in period $t-1$ regarding current values in t:

$$\pi_t = \frac{\psi\bar{\lambda}}{1-\bar{\lambda}}\left(\hat{y}_t - \hat{y}_t^n\right) + \widetilde{E}_{t-1}\left[\pi_t + \psi\Delta\left(\hat{y}_t - \hat{y}_t^n\right)\right] + \frac{\bar{\lambda}}{1-\bar{\lambda}}e_t + \bar{\lambda}\sum_{j=0}^{\infty}\left(1-\bar{\lambda}\right)^j E_{t-1-j}\Delta e_t,$$

$$\pi_t = \frac{\psi\bar{\lambda}}{1-\bar{\lambda}}\left(\hat{y}_t - \hat{y}_t^n\right) + \lambda_{t-1}E_{t-1}\left[\pi_t + \psi\Delta\left(\hat{y}_t - \hat{y}_t^n\right)\right] + (1-\lambda_{t-1})\bar{\lambda}\sum_{j=0}^{\infty}\left(1-\bar{\lambda}\right)^j E_{t-2-j}\left[\pi_t + \psi\Delta\left(\hat{y}_t - \hat{y}_t^n\right)\right] + \frac{\bar{\lambda}}{1-\bar{\lambda}}e_t + \bar{\lambda}\sum_{j=0}^{\infty}\left(1-\bar{\lambda}\right)^j E_{t-1-j}\Delta e_t. \quad (8)$$

Note that we assume heterogeneous expectations only with respect to inflation and the output gap, and not with respect to changes in the cost-push shock e_t.

2.3. Monetary Policy Rule

As usual in the New Keynesian DSGE framework, our model is closed by defining a monetary policy rule. In line with the sticky information DSGE model in Mankiw and Reis (2007), we assume a Taylor rule with interest rate smoothing, targeting actual values of inflation and the output gap:

$$\hat{i}_t = \mu_i\hat{i}_{t-1} + (1-\mu_i)\left[\mu_\pi\pi_t + \mu_{y^{\text{gap}}}\left(\hat{y}_t - \hat{y}_t^n\right)\right] + \eta_t, \quad (9)$$

where η_t is an i.i.d. shock to monetary policy. Note that we assume the central bank to perfectly perceive current values of inflation and the output gap, while at the same time accounting for the heterogeneity in agents' expectations implicitly contained in realized values of inflation and the output gap.

2.4. Endogenous Inattentiveness as a Discrete Choice under Rational Inattention

Finally, the choice between forecasts under full or sticky information is modeled as an optimization problem under rational inattention. As defined in Sims (2003), the notion of rational inattention implies that the true values of the options can be investigated, but because of agents' limited information processing capacity, it is too costly to know them with certainty. In our model, this means that agents have an idea of the relative performance of the two predictors, but are not sure about the true values. This leads to some noise in the decision process and, thus, to probabilistic choices.[8] We argue that the choice between forecasts under full

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

or sticky information can be regarded as symmetric a priori, because although forecasts under full information will generally be more accurate than those under sticky information, the costs for the full information forecast may render the sticky information forecast relatively more attractive if the difference in forecast accuracy is sufficiently small. Hence, agents do not know a priori which predictor will be more attractive in a given period, so that their prior probability of choosing each predictor without any processing of information is exactly 1/2. Therefore, we apply the framework in Matějka and McKay (2011) to derive the probabilistic choice between forecasts as the outcome to an optimization problem under rational inattention with discrete and symmetric options.[9]

We assume that predictor attractiveness of forecasts under full or sticky information is negatively affected by their past mean squared forecast errors. In the case of the full information predictor, its attractiveness is additionally reduced by the fixed cost C^{FI}, which has to be paid every period because of the recursive structure of the model. Measuring forecast performance by its mean squared forecast error is common in the literature on heterogeneous expectations; see for instance Branch (2007), De Grauwe (2010), and Branch and Evans (2011).[10]

The attractiveness of a predictor i, V^i, is thus given by the following expressions for the full information and the sticky information predictor, respectively:

$$V_t^{\mathrm{FI}} = -\left[(\hat{x}_{t-1} - E_{t-2}\hat{x}_{t-1})^2 + C^{\mathrm{FI}}\right], \tag{10}$$

$$V_t^{\mathrm{SI}} = -\left[\hat{x}_{t-1} - \bar{\lambda}\sum_{j=0}^{\infty}(1-\bar{\lambda})^j E_{t-j-3}\hat{x}_{t-1}\right]^2, \tag{11}$$

where $\hat{x}$ denotes the variable being forecasted. Under rational inattention, we assume that agents cannot fully perceive the true values of $V = (V^{\mathrm{FI}}, V^{\mathrm{SI}})$, but instead have some prior knowledge of the predictors' attractiveness given by the joint pdf $g(V)$. Moreover, agents receive signals on the choices i, where $I(V;i)$ denotes the so-called mutual information defined as the reduction in entropy of V due to the information in the signal on option i. With N discrete options, we thus get a collection $\{f(V|i)\}_{i=1}^N$, where $f(V|i)$ is the conditional posterior on V given choice i. Note that in our model $N = 2$. Finally, agents' information processing capacity is denoted by κ. We assume that κ is given exogenously and limits the amount of information processed in the mutual information given choice i, $I(V;i)$.

The optimization problem under rational inattention can thus be stated as follows. Agents choose a collection $\{f(V|i)\}_{i=1}^N$ to maximize the expected value of predictor attractiveness V subject to their information processing constraint. Because agents may switch between forecasts each period, the problem is stated as a static optimization problem, omitting time indices:

$$\max_{\{f(V|i)\}_{i=1}^N} \sum_i \int_V V^i f(V|i)dV, \tag{12}$$

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

s.t.

$$I(V;i) \leq \kappa \tag{13}$$

$$\sum_i f(V|i) = g(V), \tag{14}$$

where the last constraint imposes that the conditional posteriors are consistent with the prior. From the first-order conditions, the probability of choosing the forecast i, given a set of values in V, is then derived as[11]

$$P(i|V) = \frac{e^{V^i/\nu}}{\sum_j e^{V^j/\nu}}, \tag{15}$$

where ν is the Lagrange multiplier on the information constraint and can thus be interpreted as the shadow cost of information on forecast attractiveness. Denoting the probability of choosing the full information forecast in period t by λ_t, we thus get

$$\lambda_t = \frac{e^{V_t^{\mathrm{FI}}/\nu}}{e^{V_t^{\mathrm{FI}}/\nu} + e^{V_t^{\mathrm{SI}}/\nu}}. \tag{16}$$

The probability of choosing the full information forecast under rational inattention is thus given by a multinomial logit map that evaluates the relative attractiveness of the full information predictor.[12] A higher forecast performance relative to the sticky information predictor thus leads to a higher probability of agents choosing the full information forecast. The shadow cost of information on the predictor's attractiveness, ν, reduces the effect of actual values V on the selection probability λ_t: With high ν, information on forecast attractiveness is very costly, so that agents will base their decision more on their priors and less on actual values V. Conversely, as $\nu \to 0$, information on forecast attractiveness becomes costless and agents will choose the best forecast with probability one.

Finally, because in our model agents form expectations regarding output and inflation, we define two choice mechanisms for the probability of choosing the fully informed output and the inflation forecast, respectively. Even though both inflation and output are affected by the same shocks driving the economy, by allowing separate choice mechanisms we can account for the different coefficients attached to each shock in the dynamic processes for inflation and output. Hence, although the more general case would be a model with one choice mechanism for expectations regarding all macroeconomic variables, we argue that separate choices allow a more detailed distinction between the variables' dynamics and the resulting predictor attractiveness. Moreover, they provide a way to analyze feedback effects between the chosen degree of inattentiveness and the macroeconomic outcome, as well as the response to macroeconomic shocks separately for output and inflation.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

We thus get for the time-varying degree of attentiveness to output, λ_t^y, and inflation, λ_t^π,

$$\lambda_t^y = \frac{e^{V_{y,t}^{\mathrm{FI}}/\nu}}{e^{V_{y,t}^{\mathrm{FI}}/\nu} + e^{V_{y,t}^{\mathrm{SI}}/\nu}}, \tag{17}$$

$$\lambda_t^\pi = \frac{e^{V_{\pi,t}^{\mathrm{FI}}/\nu}}{e^{V_{\pi,t}^{\mathrm{FI}}/\nu} + e^{V_{\pi,t}^{\mathrm{SI}}/\nu}}, \tag{18}$$

where $V_{y,t}$ and $V_{\pi,t}$ are the measures of forecast attractiveness with respect to output and inflation, respectively.

2.5. Solution

For the solution of the model, we account for the infinite sum of lagged expectations in the sticky information predictor with the algorithm presented by Meyer-Gohde (2010). The author accounts for the infinite sum of lagged expectations by calculating matrices of limiting coefficients as the sum approaches minus infinity. The algorithm then solves for the undetermined coefficients of the MA(∞) representation with a generalized Schur decomposition, using the limiting coefficients as an approximation for the infinite horizon of lagged expectations. With the algorithm of Meyer-Gohde (2010), we thus solve for the equilibrium numerically.[13]

Because the degree of inattentiveness enters the model's equilibrium conditions as a time-varying coefficient, we follow Brock and Hommes (1997) and solve the model recursively over time, yielding simulated equilibrium time paths for aggregate variables. Specifically, the timing of events is as follows: Starting from an initial simulation, at the beginning of period t agents evaluate the performance of their predictor in period $t-1$ under rational inattention, and the aggregate probability of choosing the full information predictor in t is determined via the multinomial logit maps given in equations (17) and (18). Given the new values of λ_t^y and λ_t^π, the model is solved and the equilibrium MA(∞) representation is found. With the new solution and an exogenous vector of shocks, the model is simulated for period t. Again, agents evaluate their predictor's performance and probabilistically decide on their predictor, thus defining the aggregate degree of inattentiveness λ_{t+1}^y and λ_{t+1}^π. These feed back into the model solution for period $t+1$, and so forth.

As discussed in Brock and Hommes (1997) in their model with switching between a rational and an adaptive predictor, complicated equilibrium dynamics with chaotic behavior may arise under certain parameterizations with a sufficiently high intensity of choice. We thus check for the stability of the steady state for a range of initial shocks with standard deviations $-0.03 \leq \tau_y \leq 0.03$ and $-0.03 \leq \tau_\pi \leq 0.03$, as well as a range of Taylor rule coefficients $0 \leq \mu_{y^{\mathrm{gap}}} \leq 2$ and $1 < \mu_\pi \leq 2$ (keeping $1/\nu$ constant at the baseline value) and a range of parameters for the inverse of the shadow cost of information on predictor attractiveness $0 \leq 1/\nu \leq 40{,}000$ (keeping μ_π and $\mu_{y^{\mathrm{gap}}}$ constant at the baseline value). Across these

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

TABLE 1. Parameter values for calibration

Parameter		Baseline model: McCallum (2001)
Intertemporal elasticity of substitution	$1/\sigma$	0.4
Coefficient of relative risk aversion	σ	2.5
Inverse of Frisch elasticity of labour supply	η	1
Elasticity of substitution between goods	θ	35
Coefficient on output gap from PC	$\psi = \frac{\sigma+\eta}{1+\eta\theta}$	0.097
Weight of inflation target Taylor rule	μ_π	1.5
Weight of output Taylor rule	$\mu_{y^{\text{gap}}}$	0.5
Interest smoothing Taylor rule	μ_i	0.8
AR term technology shock	α_{tech}	0.950
Std IS shock	τ_y	0.015
Std PC shock	τ_π	0.015
Std MP shock	τ_i	0.01
Std technology shock	τ_{tech}	0.007
Calibration of switching parameters		
Initial degree of attentiveness to output	λ^y_{init}	0.5
Initial degree of attentiveness to inflation	$\lambda^\pi_{\text{init}}$	0.5
Average past attentiveness	$\overline{\lambda}$	0.2
Inverse of shadow cost to attention	$1/\nu$	4000
Information cost w.r.t. output	C^{FI}_y	0.001
Information cost w.r.t. inflation	C^{FI}_π	0.0001

ranges, the model always converges toward its steady state with $\hat{y} = \hat{\pi} = \hat{i} = 0$ (measured in deviations from the steady state). It thus seems that dynamics dies out quickly if the system is no longer subjected to exogenous shocks.[14]

3. RESULTS

This section presents results from numerical simulations of the model with recursive and endogenous inattentiveness. We define the model as quarterly and simulate over 1,500 periods, where the first 500 periods initialize the model and produce lagged expectations and are dropped consequently. Calibration parameters of the model correspond closely to the calibration in McCallum (2001) and are given in Table 1. Because the parameters η and θ are not included in the McCallum (2001) model, we set values to obtain a value for ψ similar to the one estimated in Mankiw and Reis (2007). The average past probability of choosing the full information predictor, $\overline{\lambda}$, is calculated from households' microdata on inflation expectations in the University of Michigan Survey of Consumers, shown in Figure 1. We find a mean value of 0.4 over six months, which gives a value of $\overline{\lambda} = 0.2$ for a quarterly model. Note that this is close to the value set for the exogenous probability of updating in Mankiw and Reis (2002) ($\lambda = 0.25$).[15] Furthermore, we set inital values of $\lambda_t^{y,\pi}$ to 0.5.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

In a recent approach, Cornea et al. (2012) estimate the inverse of the shadow cost of attention to predictor accuracy, or the intensity of choice, $1/\nu$, for a model with backward- and forward-looking expectations regarding inflation for the United States. The authors report a value of $1/\nu \approx 4$, where the attractiveness of predictors is measured by their relative absolute forecast errors (whose order of magnitude is 10^{-1}). Note that the order of magnitude of $1/\nu$ varies with V^i. In our model with squared forecast errors, V^i regarding different predictors for inflation and output varies in the range from 10^{-5} to 10^{-3}. We thus apply the mean order of magnitude 10^{-4} and set $1/\nu = 4{,}000$ in the calibration.[16]

Regarding the costs of obtaining the full information forecast C^{FI} in equation (10), we aim at finding values for C^{FI} that minimize the absolute difference between the average attractiveness measures V^{FI} and V^{SI} in the case of inflation and output expectations. This is important because our model assumes that expectations under full or sticky information are equally attractive a priori. Additionally, we want to keep the calibration of the model as parsimonious as possible, i.e., without imposing any structural difference in the choice of predictors for inflation and output. We run a grid search over the values $0 < C^{\mathrm{FI}} \leq 0.004$ for both output and inflation expectations to determine the costs C_y^{FI} and C_π^{FI}. $|V^{\mathrm{FI}} - V^{\mathrm{SI}}|$ is minimized for $C_y^{\mathrm{FI}} = 0.001$ for output expectations and $C_\pi^{\mathrm{FI}} = 0.0001$ for inflation expectations, which we use as parameters for our calibration. Note that the higher cost for full information regarding output is due to the larger variance of output relative to inflation in our model, which results in larger average forecast errors under sticky information relative to errors under full information.[17] For both inflation and output expectations the mean degree of attentiveness, $\lambda_t^{y,\pi}$, decreases as C^{FI} rises and vice versa.

To keep the model as parsimonious as possible, we set the standard deviation of demand and cost-push shocks on output and inflation equal at $\tau_y = \tau_\pi = 0.015$ percentage points. Additionally, we assume no autocorrelation in the shocks and do not include any lagged endogenous variables, except for the technology shock on natural output $\hat{y}_t^n$ and the interest-rate smoothing parameter in the Taylor rule.

3.1. Endogenous Persistence and Higher Moments of Aggregate Variables

First, we analyze simulated equilibrium time paths of output, inflation, and the nominal interest rate. As described in Section 2.5, these should be understood as dynamic equilibria, which are computed recursively over time as the equilibrium solutions of the model economy, given exogenous shocks and time-varying degrees of inattentiveness.

In Table 2, we compare the summary statistics of simulated aggregate variables from our model with those from simulations of a model with a fixed degree of inattentiveness $\overline{\lambda} = 0.2$, and of a standard DSGE model with sticky prices and full information rational expectations.[18] The sticky price DSGE is simulated with the calibration in McCallum (2001), which gives a degree of price stickiness equal to $\kappa = 0.03$. To facilitate comparison of the results, all three models are

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

TABLE 2. Sample statistics

Variable	Model	Mean	Std.	Skewness	Kurtosis	AR(1) coeff.
$\hat{y}_t$	U.S. data	1.490×10^{-15}	0.022	−0.216	3.423	0.780
	Recursive inattentiveness	1.400×10^{-3}	0.135	0.008	3.232	0.985
	Sticky information, fixed $\bar{\lambda}$	-8.590×10^{-6}	0.019	0.003	2.983	0.352
	Sticky prices	1.323×10^{-6}	0.024	0.003	2.980	0.508
π_t	U.S. data	2.770×10^{-4}	0.012	0.935	7.583	0.934
	Recursive inattentiveness	-4.989×10^{-5}	0.007	0.003	3.039	0.621
	Sticky information, fixed $\bar{\lambda}$	-3.209×10^{-6}	0.004	0.001	2.989	0.014
	Sticky prices	-1.350×10^{-5}	0.015	-9.557×10^{-5}	2.990	−0.013
$\hat{i}_t$	U.S. data	-1.770×10^{-5}	0.009	0.159	6.594	0.885
	Recursive inattentiveness	-2.605×10^{-4}	0.033	0.003	2.318	0.948
	Sticky information, fixed $\bar{\lambda}$	9.452×10^{-6}	0.015	0.004	2.971	0.750
	Sticky prices	-7.263×10^{-7}	0.013	0.001	2.978	0.684

Note: Mean values from simulating the models 1,000 times over 1,000 periods. U.S. data are analyzed for the sample period 1950q1–2011q4 for real GDP growth and for 1950m1–2011m12 for CPI inflation and short-run interest rates (3-month Treasury Bill Rates). All data are obtained from the FRED database of the Federal Reserve Bank of St. Louis and are detrended with an HP-filter. The sticky information model without recursive inattentiveness is simulated with a fixed $\bar{\lambda} = 0.2$ for both output and inflation expectations. The model with sticky prices assumes rational full information expectations and sticky prices and is calibrated with the paramters in McCallum (2001).

simulated with the same shock processes, as shown in Table 1. Hence, we do not include any autocorrelated shocks, except for the technology shock, or any lagged endogenous variables, except for the interest-rate smoothing term in the Taylor rule. The benchmark for all model simulations is given by U.S. data, which are obtained for the period 1950–2011 from the FRED database of the Federal Reserve Bank of St. Louis. We use data for quarterly real GDP growth, monthly CPI inflation and monthly 3-Months Treasury Bill rates. Furthermore, all time series are detrended with an HP filter.

Table 2 shows the mean, the standard deviation, as well as higher moments, and the degree of persistence of aggregate output, inflation, and nominal interest rates. Overall, our model with recursive inattentiveness succeeds best in matching the higher moments, as well as the persistence of U.S. data, compared with both the model with a fixed $\bar{\lambda}$ and the sticky price model. Although the mean values of both the detrended data and the model simulations are all very close to zero, the standard deviation of the data is usually matched best by either the fixed sticky information or the sticky price model. However, the higher moments of the data generated by our model with recursive inattentiveness are generally closest to those derived from U.S. data.

Regarding the degree of persistence, our model with recursive inattentiveness generates strong endogenous persistence in output, inflation, and nominal interest rates, which matches the empirical persistence from U.S. data relatively well. This is to some extent driven by the autocorrelation in the technology shock,

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

which primarily affects aggregate output, and the interest-rate smoothing parameter affecting interest rates. However, the feedback from the recursive degree of inattentiveness to the model evolution adds additional persistence to the model, as agents' probability of choosing the sticky information predictor is derived endogenously and may thus be higher than in the model with fixed sticky information. Indeed, the average degree of inattentiveness to inflation is found to be considerably higher with recursive inattentiveness than the fixed degree of inattentiveness $\overline{\lambda}$ (see Table 3 in the next section). Moreover, if the recursive degree of inattentiveness $\lambda_t^{y,\pi}$ also evolves with some persistence, i.e., agents on the average keep predictors for some periods before switching again, this will induce additional persistence into the model.

Thus, neither the fixed sticky information model nor the sticky price model with the same parameterization achieves a matching degree of persistence. This is especially true in the case of simulated inflation, where only our model with recursive inattentiveness yields a high degree of persistence, in line with the empirical evidence discussed for instance in Fuhrer and Moore (1995) and Gordon (1997).[19]

3.2. Recursive Inattentiveness

After the analysis of simulated aggregate variables in the preceding section, we turn to evaluating the recursively simulated time-varying degree of inattentiveness. Because the probability of choosing the full information forecast, $\lambda_t^{y,\pi}$, is driven by the squared forecast error of the full information versus the sticky information predictor, we first present squared forecast errors from our recursive simulation in Table 3. As expected, mean squared forecast errors from the sticky information predictor are higher than errors from the full information predictor and have a larger variation. Moreover, MSEs from the sticky information predictor are significantly more persistent, thus indicating that the forecast is misspecified.

Next, we analyze the sample statistics of recursive attentiveness to output and inflation, λ_t^y and λ_t^π, in Figure 2 and Table 4.

The aggregate degree of attentiveness to output deviates over time between one and close to zero. On the average, agents in the model use the full information predictor with a probability of about 25% for an average of 2.8 quarters before switching again.[20]

Regarding the degree of attentiveness to inflation, on the average agents forecast inflation with full information with a probability of about 40%, with the degree of attentiveness varying between 35% and 51%. Although the average attentiveness to inflation implied by the model is higher than found empirically in U.S. microdata, the range of attentiveness (± 0.1 from the mean) fits the variation of the empirical degree of attentiveness to inflation from U.S. households in Figure 1 relatively well. Although the simulated time series shown in Figure 2 is more volatile than the empirical degree of attentiveness, a cyclical pattern nevertheless emerges. Moreover, the simulated mean value for λ_t^π fits well with the empirical estimate of

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

TABLE 3. Mean squared forecast errors with recursive inattentiveness

MSE of variable	Model	Mean	Min	Max	Std.	AR(1) coeff.
π_t	Sticky information predictor	1.658×10^{-5}	0.0000	2.039×10^{-4}	2.364×10^{-5}	7.000×10^{-3}
	Full information predictor	1.414×10^{-5}	0.0000	1.686×10^{-4}	1.998×10^{-5}	-3.141×10^{-4}
	Overall recursive inattentiveness	1.510×10^{-5}	0.0000	1.807×10^{-4}	2.135×10^{-5}	-3.573×10^{-4}
$\hat{y}_t$	Sticky information predictor	0.0014	0.0000	0.0239	0.0028	0.8066
	Full information predictor	0.0003	0.0000	0.0003	0.0004	0.0745
	Overall recursive inattentiveness	0.0004	0.0000	0.0064	0.0006	−0.0258

Note: Mean values from simulating 1,000 times over 1,000 periods.

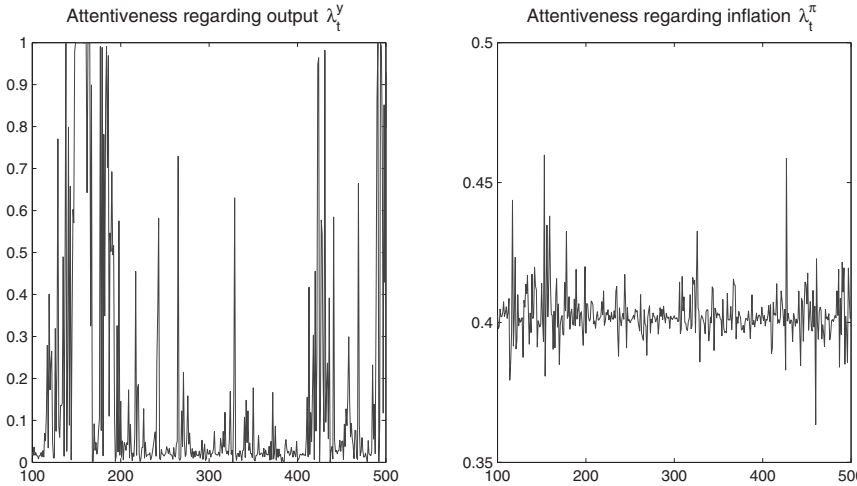


FIGURE 2. Recursively simulated attentiveness over time.

TABLE 4. Recursive inattentiveness

Variable	λ_t^y	λ_t^π
Mean	0.251	0.404
Min	6.777×10^{-5}	0.352
Max	1.000	0.508
Std.	0.370	0.011
Average cycle length	2.800	1.443

Note: Mean values from simulating 1,000 times over 1,000 periods. The average cycle length is calculated in quarters.

$\lambda = 0.408$ derived from a sticky information Phillips curve for the United States in Kiley (2007).

Overall, the simulation results suggest that agents are on the average more attentive to inflation than to output and switch inflation predictors more frequently than their output predictors. Moreover, the degree of attentiveness to output is much more volatile and includes periods when all agents are attentive to recent output developments and vice versa. This may be explained both by the significantly higher variance of the simulated output series than of simulated inflation and by the higher persistence of simulated output than of inflation (see Table 2): Because of the relatively higher variance of output, the attractiveness measures for output expectations, V_y, vary relatively more than V_π for inflation expectations, resulting in a wider range for λ_y than for λ_π. In contrast, the higher persistence of simulated output compared to inflation implies that outdated information regarding output generally gives better forecasts than in the case of inflation, which results in higher average attentiveness to inflation and more frequent switching of inflation

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

TABLE 5. Recursive inattentiveness and macroeconomic variables

Correlation with	λ_t^y	λ_t^π
Level π_t	−0.0049	−0.0014
Variance π_t	0.1225	0.1730
Level $\hat{y}_t$	0.0078	0.0027
Variance $\hat{y}_t$	0.3716	0.0983
Level $\hat{i}_t$	−0.0062	−0.0035
Variance $\hat{i}_t$	−0.0227	0.0039

Note: Mean values from simulating 1,000 times over 1,000 periods.

predictors. Nevertheless, the simulated time series for λ_t^y and λ_t^π in Figure 2 suggest that the volatility of attention to both output and inflation is positively correlated: Over time, periods of strong movements in attentiveness to both variables are followed by periods of relative stability in the degree of attentiveness, and vice versa.

Next, we analyze the relation between agents' choice of predictors and the macroeconomic conditions in the model economy. Table 5 presents correlation coefficients of λ_t^y and λ_t^π with the level and the time-varying variance of aggregate output, inflation and nominal interest rates. We find that the degree of attentiveness is positively correlated with the variance of the variable to be forecast. Thus, agents are increasingly likely to opt for costly expectations with full information as the variability of the forecast variable and, hence, the uncertainty surrounding its forecast rises and remain inattentive to new developments in the variable otherwise. This is in line with predictions from models of rational inattention, as in Sims (2003), where optimal attention is positively affected by the variance of a variable, because this gives a signal of the shocks driving the variable's process [see also Mackowiak and Wiederholt (2009)]. Our model's predictions thus confirm the link between sticky information and rational inattention models. Furthermore, we find some feedback between attention to output and inflation, because a higher variability of output is also positively correlated with the degree of attentiveness to inflation and vice versa.

3.3. Monetary Policy with Recursive Inattentiveness

After evaluating the statistical properties of simulated series for aggregate variables and recursive inattentiveness, we turn to analyzing the effects of a monetary policy shock. Specifically, we are interested in whether the model with recursive inattentiveness can generate the hump-shaped responses of both output and inflation after a monetary policy shock present in the sticky information models of Mankiw and Reis (2002, 2006) and found empirically, for instance, in Rotemberg and Woodford (1997) and Christiano et al. (2005). Because our model has a dynamic equilibrium with time-varying parameters λ_t^y and λ_t^π, however, overall impulse

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

responses in terms of the MA(∞) coefficients cannot be derived, because the MA representation of the model changes each period. Therefore, over 200 simulation periods, we plot impulse responses of both output and inflation to a monetary policy shock with time-varying degrees of attentiveness λ_t^y and λ_t^π, shown in Figure 3.

From Figure 3 we see that over the simulation period, impulse responses of inflation to the monetary policy shock generally show a hump-shaped pattern with some overshooting, whereas responses of output are persistent, albeit not hump-shaped. Moreover, in some periods, both variables show larger and peaked impulse responses. These peaked impulse responses occur whenever the degree of attentiveness is very high; i.e., agents are very likely to choose the full information forecast and thus incorporate the shock more quickly into their expectations. In contrast, when the degree of attentiveness is relatively low, news about the monetary policy shock is observed and processed more slowly, yielding a delayed impulse response. Because λ_t^y is relatively more volatile than λ_t^π, periods with peaked impulse responses are mostly periods with a high degree of attentiveness to recent output developments.

In order to show the effects of high versus low degrees of attentiveness on the responses to a monetary policy shock more clearly, in Figure 4 we present mean impulse responses of output and inflation for periods when both λ_t^y and λ_t^π are in the top or bottom twentieth percentile. Again, we see that impulse responses with a high degree of attentiveness, i.e., with values of $\lambda_t^{y,\pi}$ in the top 20%, are stronger on impact and less persistent than those in periods with a high degree of inattentiveness ($\lambda_t^{y,\pi}$ in the bottom 20%). The effect is more pronounced for impulse responses of inflation, where we observe a hump-shaped response to the monetary policy shock only in periods with a high degree of inattentiveness. In contrast, in periods with a low degree of inattentiveness, inflation shows a strong negative response on impact, which overshoots in the second period and then coverges toward zero.

4. CONCLUSION

We present a sticky information DSGE model with recursive inattentiveness and heterogeneous expectations. Extending the models of sticky information developed in Mankiw and Reis (2002, 2007), we endogenize the aggregate degree of inattentiveness as the outcome to a probabilistic choice between a costly full information predictor and a costless sticky information predictor. Assuming that agents obtain only the forecast, but not the method, and have only limited knowledge of the relative performance of each predictor, the choice between predictors can be derived as a discrete choice under rational inattention; see Matějka and McKay (2011). Thus, our model provides a link between models of sticky information and models of rational inattention. The optimal choice of predictors under rational inattention yields a multinomial logit map for the probability of choosing the full information predictor, which is also frequently applied in the literature on heterogeneous expectations (Brock and Hommes, 1997). Recursive simulation of the

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

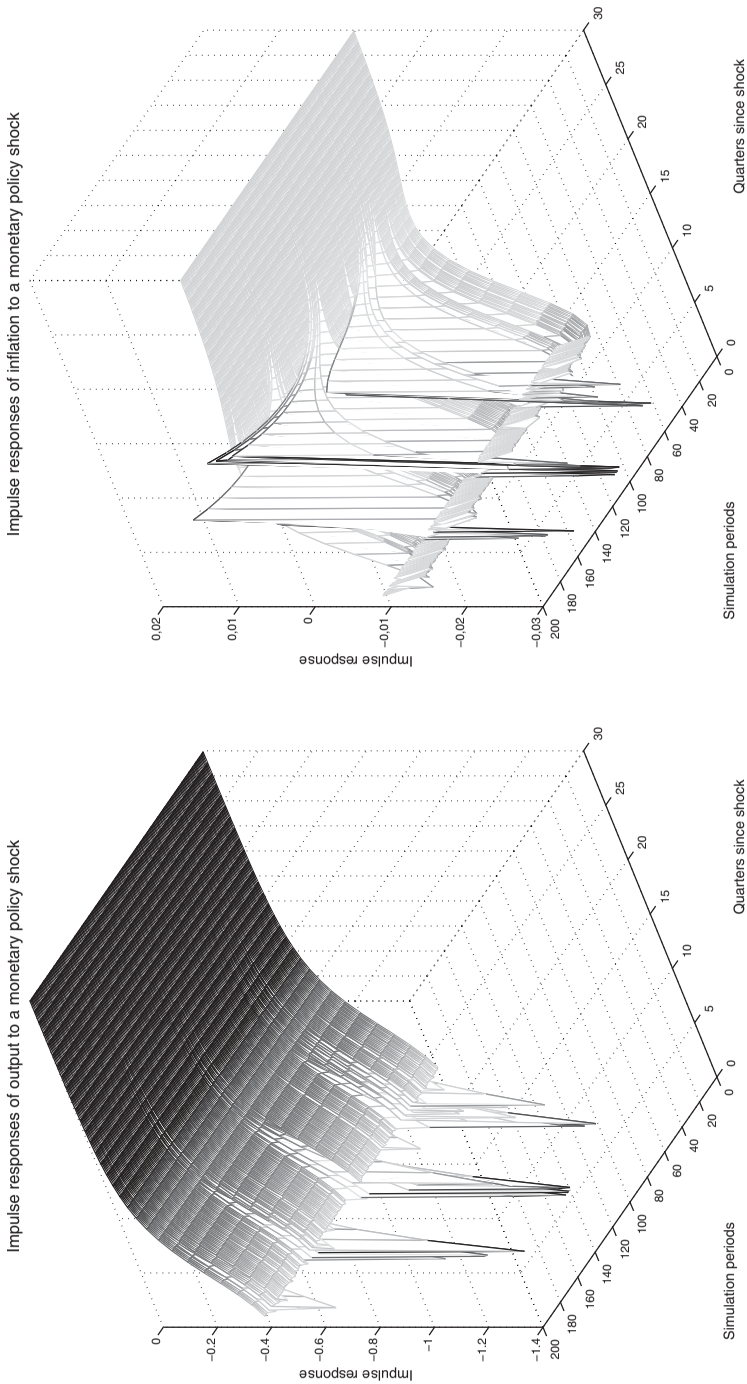


FIGURE 3. Simulated responses to a monetary policy shock.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

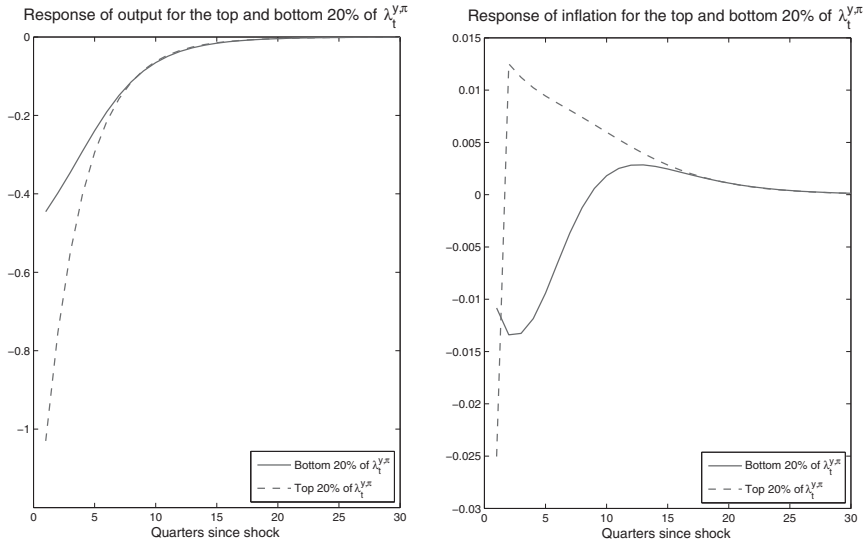


FIGURE 4. Mean impulse responses to a monetary policy shock under high/low degrees of attentiveness.

model then yields a dynamic equilibrium path with endogenous and time-varying degrees of inattentiveness to both output and inflation.

We compare the moments and the persistence of recursively simulated paths for aggregate output, inflation, and nominal interest rates from our model with those from simulations of a model with sticky information and a fixed degree of inattentiveness and a model with sticky prices. Our model with recursive inattentiveness generally succeeds best in matching the higher moments and the persistence of U.S. data. It thus seems that accounting for recursively changing degrees of inattentiveness yields more realistic simulations of the model economy than models with fixed inattentiveness or full information rational expectations. Especially in the case of inflation, neither the model with fixed sticky information nor the model with sticky prices are able to reproduce the stylized fact of high inflation persistence [see Fuhrer and Moore (1995)].

Simulations with recursive inattentiveness suggest that on the average, attentiveness to inflation is higher than attentiveness to output, and inflation predictors are switched more frequently. This may be explained by the relatively higher persistence of simulated output compared to inflation: As output changes more sluggishly, on the average outdated information on output generates better forecasts than in the case of inflation. Moreover, the probability of choosing the full information predictor is positively correlated with the variance of the forecasted variable for both output and inflation expectations. Agents in our model thus behave as in the rational inattention models in Sims (2003) and Mackowiak and Wiederholt (2009), because they are more likely to pay attention to recent

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

developments of output and inflation in times of high volatility in the economy and ignore smaller changes.

With respect to the stylized fact of a hump-shaped response to a monetary policy shock, we find that our model with recursive inattentiveness can reproduce a hump-shaped impulse response of inflation when the degree of inattentiveness is sufficiently high, whereas the response of output is persistent, but not hump-shaped. Generally, the degree of persistence in the response to a monetary policy shock depends on the degree of inattentiveness chosen, where higher attentiveness implies that shocks are accomodated more quickly.

Overall, by allowing a choice between a full information predictor and a sticky information predictor, we can thus analyze how aggregate attentiveness to new information is distributed between output and inflation, how it is affected by changes in the level and variance of aggregate variables, and how changes in the degree of inattentiveness in turn affect the macroeconomy. Our results may provide useful insights for economic policy, as they suggest that the evolution of aggregate variables over time and their response to economic shocks in general show strong persistence. However, this may change during periods of high volatility in the economy, as this induces a higher degree of attentiveness by economic agents and, hence, quicker responses to economic shocks.

NOTES

1. Empirical estimates of the share of agents updating information each period are given for the United States in Carroll (2003) and for a European sample in Döpke et al. (2008a, 2008b).

2. Recently, a number of papers have studied the derivation of aggregate relations with heterogeneous expectations from the micro foundations in a New Keynesian framework. Massaro (2013) derives reduced-form relations and stresses that the Taylor principle may no longer be sufficient for determinacy under certain types of heterogeneous expectations. Woodford (2013) evaluates the implications of different types of expectations for well-known results such as the policy trade-off between inflation and output stabilization. Finally, Kurz et al. (2013) develop axioms regarding the nature of beliefs necessary for the aggregation of heterogeneous expectations and derive aggregate relations which also depend on the market distribution of beliefs. For our model, we leave the full derivation of aggregate relations from the micro level for future research.

3. Note that $\bar{\lambda}$ may only take on values between zero and one, because otherwise the weight of the sticky information predictor in the aggregate forecast $\widehat{E}_t$ could be zero even if $\lambda_t \neq 1$.

4. Microdata for individual changes in U.S. consumers' inflation expectations from the University of Michigan Survey of Consumers show that over a period of six months between two interviews, quantitative inflation expectations are changed much more frequently than the qualitative, more general, inflation outlook. This might be due to fine tuning of quantitative expectations, but could also be indicative of a certain memory loss regarding the quantitative forecast stated in the first interview; see Dräger and Lamla (2012).

5. The model presented here does not allow for an endogenous degree of agents' memory with respect to forecasts obtained in previous periods. We thus leave this issue for future research.

6. Note that this assumption is similiar to those necessary in Reis (2006b) for the derivation of the sticky information Phillips curve and is commonly applied in the literature on heterogeneous expectations. For an analysis of optimal monetary policy in models with imperfect information processing accounting for higher-order beliefs, see Adam (2007, 2009).

7. We define u_t as an i.i.d. demand shock capturing exogenous changes in aggregate demand components other than consumption.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

8. Note that this line of argument is similar to the derivation of probabilistic choices in random utility models; see for instance Hommes (2006).

9. The detailed solution of the optimization problem follows the analysis in Matějka and McKay (2011) and is given in the Appendix.

10. Similarly, Branch et al. (2009) endogenize the degree of inattentiveness of firms by proposing a loss function that measures expected profit losses as the squared difference of a firm's individual price from the optimal price, given some economywide degree of inattentiveness $\overline{\lambda}$. Brock and Hommes (1997) show that using mean squared forecast errors or realized net profits as the measure of predictor fitness yields similar results.

11. The derivation of equation (15) is shown in the Appendix.

12. Although we have derived the multinomial logit map in (16) from an optimization problem under rational inattention, the multinomial logit model is also frequently applied in the literature on discrete choice theory, where it is mostly derived from a random utility model [see Manski and McFadden (1981)]. Note that the so-called intensity of choice, which governs the relative importance of the deterministic and the random part of utility, is inversely related to the shadow cost of information on forecast accuracy, ν, in our derivation. A New Keynesian model with heterogeneous expectations and switching between a costly full information and a costless sticky information predictor, modeled as in Brock and Hommes (1997), would thus give a mathematically equivalent model.

13. MATLAB codes for the Meyer-Gohde (2010) solution algorithm are freely available at http://ideas.repec.org/c/dge/qmrbcd/171.html.

14. Simulation results are available from the author upon request.

15. We checked for robustness of our results to alternative values of the weighting parameter $\overline{\lambda}$ in the range $0 < \overline{\lambda} < 1$. As $\overline{\lambda}$ increases, past expectations receive a lower weight in the sticky information predictor. Although our main results remain robust, a higher parameter $\overline{\lambda}$ induces more variation in the degree of attentiveness, especially with respect to inflation (λ_t^{π}). This leads to a higher variance of the simulated macro data, with reduced persistence, especially in simulated inflation.

16. Note that a relatively low shadow cost of attention implies that agents have a relatively low uncertainty regarding the attractiveness of the predictors. In that sense, our calibration may be regarded as conservative with respect to the degree of persistence the model could yield because of uncertainty about predictor attractiveness. We checked for robustness of our results with values of $0 < 1/\nu \leq 40{,}000$. Generally, a higher shadow cost of attention results in less switching between predictors, whereas lower costs induce larger variations of the degree of inattentiveness, $\lambda_t^{\pi,y}$. Our results remain robust with changes in $1/\nu$. As would be expected, with a lower shadow cost of attention to predictor accuracy, the model yields a higher correlation of $\lambda_t^{\pi,y}$ with the variance of the forecasted variable and somewhat more persistent macro variables. We also checked for robustness of our results when we assign individual shadow costs to λ_t^{π} and λ_t^{y} to exactly match the order of magnitude of the fitness measures. Again, this affects the evolution of $\lambda_t^{\pi,y}$ somewhat, but does not change our general results. Therefore, we opt for the more parsimonious model with the same shadow cost of attentiveness to predictor attractiveness, ν, for both predictors. Results for all robustness checks are available upon request.

17. Moreover, a higher information cost for output vs. inflation might be motivated by more frequent news on inflation due to monthly inflation data releases compared to quarterly output estimates. Additionally, higher information costs for the output gap may arise because of uncertainty regarding the natural level of output.

18. Note that we omit a comparison to simulations from rational inattention models such as the models in Mackowiak and Wiederholt (2009) and Woodford (2009), because their structure differs significantly from that of our model. Moreover, they are solved with a different solution algorithm and, thus, simulation results are not easily comparable.

19. Note that the sticky information model with a fixed $\overline{\lambda}$ in Mankiw and Reis (2002) assumes an autocorrelated process for the money supply shock driving the economy. Because, in their simple model, money supply is equal to aggregate demand, this assumption may drive the result of a strong persistence of inflation. The same holds true for sticky price DSGE models with either autocorrelated cost-push shocks or a hybrid New Keynesian Phillips curve including a lagged inflation term.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

20. The average cycle length is defined as the average number of quarters that λ_t^y or λ_t^π stay close to constant before switching occurs again. For this measure, we define λ_t^y or λ_t^π as "close to constant" if their value in the current period does not deviate from their value in the previous period by more than a threshold of 0.001.

REFERENCES

Adam, Klaus (2007) Optimal monetary policy with imperfect common knowledge. *Journal of Monetary Economics* 54, 267–301.

Adam, Klaus (2009) Monetary policy and aggregate volatility. *Journal of Monetary Economics* 56, S1–S18.

Anufriev, Mikhail, Tiziana Assenza, Cars H. Hommes, and Domenico Massaro (2013) Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics* 17, 1574–1604.

Ball, Laurence (2000) Near-Rationality and Inflation in Two Monetary Regimes. NBER working paper 7988.

Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis (2005) Monetary policy for inattentive economies. *Journal of Monetary Economics* 52, 703–725.

Berardi, Michele and John Duffy (in press) Real time, adaptive learning via parameterized expectations. *Macroeconomic Dynamics*.

Branch, William A. (2004) The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal* 114, 592–621.

Branch, William A. (2007) Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control* 31, 245–276.

Branch, William A., John Carlson, George W. Evans, and Bruce McGough (2006) Adaptive Learning, Endogenous Inattention and Changes in Monetary Policy. Federal Reserve Bank of Cleveland working paper 0610.

Branch, William A., John Carlson, George W. Evans, and Bruce McGough (2009) Monetary policy, endogenous inattention and the volatility trade-off. *Economic Journal* 119, 123–157.

Branch, William A., Troy Davig, and Bruce McGough (2013) Adaptive learning in regime-switching models. *Macroeconomic Dynamics* 17, 998–1022.

Branch, William A. and George W. Evans (2011) Monetary policy and heterogeneous expectations. *Economic Theory* 47, 365–393.

Branch, William A. and Bruce McGough (2009) A New Keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control* 33, 1036–1051.

Brazier, Alex, Richard Harrison, Mervyn King, and Tony Yates (2008) The danger of inflating expectations of macroeconomic stability: Heuristic switching in an overlapping-generations monetary model. *International Journal of Central Banking* 4, 219–254.

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

Calvo, Guillermo A. (1983) Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12, 383–398.

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

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

Clarida, Rirchard, Jordi Galí, and Mark Gertler (1999) The science of monetary policy: A New Keynesian perspective. *Journal of Economic Literature* 37, 1661–1707.

Cornea, Adriana, Cars H. Hommes, and Domenico Massaro (2012) Behavioral Heterogeneity in U.S. Inflation Dynamics. CeNDEF working paper 12-03.

De Grauwe, Paul (2010) Top-down versus bottom-up macroeconomics. *CESifo Economic Studies* 56, 465–497.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

De Grauwe, Paul (2011) Animal spirits and monetary policy. *Economic Theory* 47, 423–457.
Döpke, Jörg, Jonas Dovern, Ulrich Fritsche, and Jiri Slacalek (2008a) Sticky information Phillips curves: European evidence. *Journal of Money, Credit and Banking* 40, 1513–1520.
Döpke, Jörg, Jonas Dovern, Ulrich Fritsche, and Jiri Slacalek (2008b) The dynamics of European inflation expectations. *B.E. Journal of Macroeconomics* 8, Article 12.
Dräger, Lena and Michael J. Lamla (2012) Updating inflation expectations: Evidence from micro-data. *Economics Letters* 117, 807–810.
Fuhrer, Jeff and George Moore (1995) Inflation persistence. *Quarterly Journal of Economics* 110, 127–159.
Gordon, Robert J. (1997) The time-varying NAIRU and its implications for economic policy. *Journal of Economic Perspectives* 11, 11–32.
Hahn, Volker (2014) Transparency in monetary policy, signaling, and heterogeneous information. *Macroeconomic Dynamics* 18, 369–394.
Hommes, Cars H. (2006) Heterogeneous agent models in economics and finance. In Leigh Tesfatsion and Kenneth L. Judd (eds.), *Handbook of Computational Economics* (1st ed.), pp. 1109–1186. Amsterdam: North-Holland.
Hommes, Cars H. (2011) The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control* 35, 1–24.
Hommes, Cars H., Joep Sonnemans, Jan Tuinstra, and Henk van de Velden (2005) Coordination of expectations in asset pricing experiments. *Review of Financial Studies* 18, 955–980.
Kiley, Michael T. (2007) A quantitative comparison of sticky-price and sticky-information models of price setting. *Journal of Money, Credit and Banking* 39, 101–125.
Kurz, Mordecai, Giulia Piccillo, and Howei Wu (2013) Modeling diverse expectations in an aggregate New Keynesian model. *Journal of Economic Dynamics & Control* 37, 1403–1433.
Lansing, Kevin J. (2009) Time-varying U.S. inflation dynamics and the New Keynesian Phillips curve. *Review of Economic Dynamics* 12, 304–326.
Mackowiak, Bartosz A. and Mirko Wiederholt (2009) Optimal sticky prices under rational inattention. *American Economic Review* 99, 769–803.
Mackowiak, Bartosz A. and Mirko Wiederholt (2010) Business Cycle Dynamics under Rational Inattention. CEPR discussion paper 7691.
Mankiw, N. Gregory and Ricardo Reis (2002) Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.
Mankiw, N. Gregory and Ricardo Reis (2003) What measure of inflation should a central bank target? *Journal of the European Economic Association* 1, 1058–1086.
Mankiw, N. Gregory and Ricardo Reis (2006) Pervasive stickiness. *American Economic Review* 96, 164–169.
Mankiw, N. Gregory and Ricardo Reis (2007) Sticky information in general equilibrium. *Journal of the European Economic Association* 5, 603–613.
Manski, Charles F. and Daniel L. McFadden (eds.) (1981) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.
Massaro, Domenico (2013) Heterogeneous expectations in monetary DSGE models. *Journal of Economic Dynamics and Control* 37, 680–692.
Matějka, Filip and Alisdair McKay (2011) Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. CERGE-EI working paper 442.
McCallum, Bennett T. (2001) Should monetary policy respond strongly to output gaps? *American Economic Review* 91, 258–262.
Meyer-Gohde, Alexander (2010) Linear rational-expectations models with lagged expectations: A synthetic method. *Journal of Economic Dynamics & Control* 34, 984–1002.
Paciello, Luigi and Mirko Wiederholt (2014) Exogenous information, endogenous information and optimal monetary policy. *Review of Economic Studies* 81, 356–388.
Pfajfar, Damjan and Blaz Zakelj (2011) Experimental Evidence on Inflation Expectation Formation. Tilburg University, CentER discussion paper 2011-091.

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

Reis, Ricardo (2006a) Inattentive consumers. *Journal of Monetary Economics* 53, 1761– 1800.
Reis, Ricardo (2006b) Inattentive producers. *Review of Economic Studies* 73, 793–821.
Rotemberg, Julio J. and Michael Woodford (1997) An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual* 12, 297–346.
Sims, Christopher A. (2003) Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690.
Thomas, Lloyd (1999) Survey measures of expected U.S. inflation. *Journal of Economic Perspectives* 13, 125–144.
Woodford, Michael (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.
Woodford, Michael (2009) Information-constrained state-dependent pricing. *Journal of Monetary Economics* 56, S100–S124.
Woodford, Michael (2013) Macroeconomic analysis without the rational expectations hypothesis. *Annual Review of Economics* 5, 303–346.

MATHEMATICAL APPENDIX: SOLUTION OF THE RATIONAL INATTENTION PROBLEM

In this section, we present the solution of the rational inattention problem with discrete choices as in Matějka and McKay (2011). The optimization problem is given by

$$\max_{\{f(V|i)\}_{i=1}^{N}} \sum_i \int_V V^i f(V|i) dV \tag{A.1}$$

s.t.

$$I(V;i) \leq \kappa \tag{A.2}$$

$$\sum_i f(V|i) = g(V), \tag{A.3}$$

where the last constraint imposes that the conditional posteriors are consistent with the prior. Following information theory, the mutual information $I(V|i)$ can be written as

$$\begin{aligned} I(V|i) &= \sum_i \int_V f(V|i) \log \frac{f(V|i)}{g(V)} dV \\ &= \sum_i \int_V f(V|i) \log \frac{f(V|i)}{\sum_j f(V|j)} dV, \end{aligned} \tag{A.4}$$

where the last equality replaces the prior with the constraint in (A.3). Substituting (A.4) into the information constraint in (A.2), we get the following Lagrangian for the optimization problem under rational inattention:

$$\mathcal{L} = \sum_i \int_V V^i f(V|i) dV - \nu \left[\sum_i \int_V f(V|i) \log \frac{f(V|i)}{\sum_j f(V|j)} dV - \kappa \right]$$

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press

$$-\int_V \mu(V)\left[\sum_i f(V|i) - g(V)\right]dV, \tag{A.5}$$

where ν is the Lagrange multiplier on the information constraint and $\mu(V)$ are Lagrange multipliers on the constraint in (A.3). The first-order condition with respect to $f(V|i)$ is then derived as follows:

$$\frac{\partial \mathcal{L}}{\partial f(V|i)} = V^i - \nu \log \frac{f(V|i)}{\sum_j f(V|j)} - \nu f(V|i)\frac{\sum_j f(V|j)}{f(V|i)}\frac{1}{\sum_j f(V|j)}$$
$$+\nu \sum_k f(V|k)\frac{\sum_j f(V|j)}{f(V|k)}\frac{f(V|k)}{\left[\sum_j f(V|j)\right]^2} - \mu(V) \stackrel{!}{=} 0. \tag{A.6}$$

Rearranging, canceling, and using $\sum_j f(V|j) = g(V)$ then gives

$$V^i - \nu \log \frac{f(V|i)}{g(V)} - \mu(V) = 0. \tag{A.7}$$

Assuming a positive shadow cost of information ($\nu > 0$), we can then solve for the conditional posterior $f(V|i)$:

$$\begin{aligned} \nu \log f(V|i) &= V^i - \mu(V) + \nu \log g(V), \\ \log f(V|i) &= V^i/\nu - \mu(V)/\nu + \log g(V), \\ f(V|i) &= e^{V^i/\nu} e^{-\mu(V)/\nu} g(V) \quad \forall i. \end{aligned} \tag{A.8}$$

Substituting for $f(V|i)$ in the constraint in (A.3) with the expression in (A.8), we get

$$\begin{aligned} \sum_j e^{V^j/\nu} e^{-\mu(V)/\nu} g(V) &= g(V), \\ e^{-\mu(V)/\nu} g(V) \sum_j e^{V^j/\nu} &= g(V), \\ e^{-\mu(V)/\nu} &= \frac{1}{\sum_j e^{V^j/\nu}}. \end{aligned} \tag{A.9}$$

This can be used in equation (A.8) to eliminate $e^{-\mu(V)/\nu}$, which gives the following first-order condition:

$$f(V|i) = \frac{g(V)e^{V^i/\nu}}{\sum_j e^{V^j/\nu}} \quad \forall i, \tag{A.10}$$

where ν can be interpreted as the shadow cost of information on forecast attractiveness. The probability of choosing the forecast i given a set of values in V is then given by $P(i|V) = f(V|i)/g(V)$, which can be reformulated using (A.10) as follows:

$$P(i|V) = \frac{e^{V^i/\nu}}{\sum_j e^{V^j/\nu}}. \tag{A.11}$$

https://doi.org/10.1017/S1365100514000741 Published online by Cambridge University Press