# Empirical calibration of adaptive learning☆

Michele Berardi[a], Jaqueson K. Galimberti[b],*

[a] *University of Manchester, United Kingdom*
[b] *ETH Zurich, Switzerland*

**ABSTRACT**

Adaptive learning introduces persistence in the evolution of agents' beliefs over time, helping explain why economies present sluggish adjustments towards equilibrium. The pace of this learning process is directly determined by the gain parameter. We document and evaluate gain calibrations for a broad range of model specifications with macroeconomic data, also developing alternative approaches to the endogenous determination of time-varying gains in real-time. Our key findings are that learning gains are higher for inflation than for output growth and interest rates, and that calibrations to match survey forecasts are lower than those derived according to forecasting performance, suggesting some degree of bounded rationality in the speed with which agents update their beliefs.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

The practice of modeling expectations through the use of adaptive learning in macroeconomic models has become increasingly popular in the recent applied literature.[1] By allowing persistence in the evolution of agents' beliefs over time,

---

[1] For some time the literature on adaptive learning developed with a focus on the theoretical debate about the convergence of learning to rational expectations equilibria (see Evans and Honkapohja, 2001). The seminal contribution in the applied literature is Sargent (1999), with a study on the role of learning by the monetary authority in the evolution of US inflation rates, followed by Marcet and Nicolini (2003), who showed how learning can help explain the recurrence of hyperinflation episodes observed in several South-American countries during the 1980's. The literature that followed then unfolded into two main strands: policy-oriented studies of particular episodes of macroeconomic turmoil (mainly the US Great Inflation period during the 1970's; see, e.g., Bullard and Eusepi, 2005; Orphanides and Williams, 2005; Primiceri, 2006; Sargent et al., 2006; Milani, 2008); and business cycles modeling studies (Williams, 2003; Huang et al., 2009; Chevillon et al., 2010; Eusepi and Preston, 2011; Slobodyan and Wouters, 2012; Milani, 2007; Milani, 2011; Milani, 2014; Ormeño and Molnár, 2015).

adaptive learning provides an alternative[2] explanation for the common observation of inertia in the dynamics of macroeconomic variables. Importantly, the persistence introduced by learning is directly related to the calibration of the learning algorithm assumed to represent the process through which agents update their beliefs. The implementation of this recursive algorithm requires the pre-specification of a sequence of learning gain values, or of a mechanism through which these gains are determined in real-time. This paper is devoted to investigate this issue empirically, particularly aiming to provide a guide for applied researchers on plausible calibration values of the learning gain.

For applied purposes, the importance of the gain is related to its role in determining the statistical properties of the evolution of agents' beliefs. While adaptive learning establishes a new channel for the dynamic dissipation of structural and expectational shocks throughout the economy, the intensity of these effects is directly determined by the value of the learning gain. On one hand, higher gains can be associated with higher degrees of variability in agents' beliefs, which could ultimately improve the explanatory power of models with learning; on the other, it is under smaller gain values that learning becomes more sluggish, hence increasing the persistence of deviations from the equilibrium path.[3]

In this context, our main contribution in this paper is the documentation of plausible gain calibrations according to their quality in forecasting macroeconomic variables and in matching survey forecasts. Moreover, we explore different elements of interest for applied research, such as: several model specifications; a proper account of real-time data restrictions; alternative foundations for an endogenous determination of the gains; and different roles for the gains. Although our results indicate that the gain calibrations are sensitive to all of these features, one of our most important findings is that the gains should be allowed to differ for different variables requiring forecasting in order to reflect their own statistical properties. We also find that calibrations to match survey forecasts are found to be lower than those derived according to the forecasting performance, suggesting the existence of some degree of bounded rationality in the speed with which agents update their beliefs.

## 1.1. Approach and main results

In order to shed some light on the gain calibration issue we develop an empirical framework that mimics a real-time learning-to-forecast process. We document renewed numerical calibrations of the gains for empirical applications with US quarterly data on inflation, output growth, and interest rates. One key feature in our analysis is our coverage of a broad range of model specifications to represent agents' beliefs: we explore all possible combinations of the variables above in vector autoregressive (VAR) forecasting models with up to four lags. This is motivated in Section 3 by the pervasive possibilities of misspecification introduced by the adaptive learning approach, particularly with respect to agents' perceived laws of motion (PLM).[4] Regarding the learning algorithm, our focus is on the least squares (LS) algorithm, which has received most of the attention in the literature.

Using such framework, detailed in Section 4, we conduct several gain calibration exercises covering data over the period from 1981 to 2012; data from earlier periods are used to initialize the learning algorithm and the calibrations. We segment the calibrations according to different assumptions in the determination of the learning gains, particularly with respect to the measure used for their selection and their variation over time. Regarding the selection measure, we distinguish between two alternatives depending on the reference data: actual-based calibrations are selected by maximizing the accuracy of the forecasts; survey-based calibration, instead, are selected by maximizing the resemblance of the learning-based forecasts to those obtained from surveys with professional forecasters.

These calibrations are documented in Section 5, where we find evidence of a great degree of heterogeneity in the gain calibrations, depending mainly on the variable forecasted and the lag length of the forecasting model. Inflation presented the highest gain calibrations, followed by output growth and interest rates, with maximum averaged values on VAR(1) models of about 0.11, 0.02, and 0.005, respectively; we associate inflation's prominence to a higher degree of non-stationarity in its determination, hence requiring a higher learning gain to track its time-varying properties. These calibrations then tended to decrease as the VAR lag order increased, which indicate that the non-stationarity captured by the higher gain values under the more restricted specifications can be due to model misspecification. Interestingly, the calibrations on the forecasting models of interest rates were found to be very small, suggesting that the real-time tracking provided by learning adds little predictive content to the estimates obtained in the pre-evaluation sample period.

Another key finding from the documented gain calibrations is that the actual-based approach leads to higher gain calibrations than the survey-based. As we discuss later, this result indicates that the professional forecasters learn at a slower

---

[2] Common extensions to add persistence in the standard framework of dynamic stochastic general equilibrium (DSGE) models include consumption habits, price indexation and adjustment costs (Christiano et al., 2005), or more directly through the introduction of persistent structural shocks (Smets and Wouters, 2007).

[3] This is consistent with a well known trade-off between speed and accuracy of recursive estimation from the engineering literature: on one extreme, tracking can be slower than the system actual time variations, but with less noisy estimates; on the other extreme, tracking can be made as rapid as the time-varying context, but with estimates much more contaminated by noise (see Benveniste et al., 1990). On the link between gain values and the degree of variability in agents' beliefs see also Carceles-Poveda and Giannitsarou (2007).

[4] It is important to note that although we are not evaluating the effects of learning within a particular structural model, our focus on multiple PLMs also allows us to obtain results that are useful for a wide range of applications of adaptive learning in macroeconomics; namely, we hope our results provide some guidance on plausible calibrations for the introduction of learning in models where agents are assumed to form expectations about the variables under analysis here.

rate than it would be optimal to forecast those macro variables, which constitutes evidence consistent with the bounded rationality view of adaptive learning. In order to assess the different behavioral assumptions on the determination of the learning gains, in Section 6 we conduct a comparative analysis on the forecasting performance of the gain calibrations. Our results point to mixed evidence, providing support to the choice of different gain calibrations, depending mainly on the variable forecasted and the model specification. But the effects of these different gain assumptions on the forecasting performance of the learning models are found to be of limited statistical relevance, particularly relative to the role played by the model specifications.

Finally, in Section 7 we evaluate the effects of further dimensions of applied interest over the gain calibrations, such as: the use of revised instead of real-time data; alternative data definitions for the variables requiring the formation of expectations; and calibrations representative of the evolution of policymakers beliefs. In what follows in the next section we begin our analysis by relating this study to the previous literature and the potential rationales for the determination of the learning gains.

## 2. Related literature

### 2.1. Theoretical foundations

From a theoretical point of view, the gain determines both whether convergence to a rational expectations (RE) equilibrium takes place and the dynamic properties of the transition towards that equilibrium. In order to obtain positive convergence results, the analysis of learning places strong restrictions on the sequence of gains. Examples of these restrictions include the requirement of a "sufficiently small" gain to guarantee weaker (in distribution) convergence results (see Evans and Honkapohja, 2001, Ch.7). Furthermore, a phenomenon known as "escape dynamics," recurrently pushing the economy away from its equilibrium, has been found to have its frequency of occurrence associated to the value of the learning gain (Cho et al., 2002).

There are different interpretations for the role played by the learning gains. Sargent and Williams (2005) and Evans et al. (2010) establish a Bayesian interpretation of learning. Other than relating the learning algorithm to the Kalman filter, the Bayesian optimal estimator of a random walk time-varying parameters model, the priors also carry implications about the specific calibrations of the learning gains. Hence, in a Bayesian framework the gains can be interpreted as representing agents' prior beliefs about the evolution of the economy.

Another useful interpretation given to the learning gains is that of determining the memory of the learning mechanism (see, e.g., Barucci, 1999; Honkapohja and Mitra, 2006). From the estimation problem originating the learning algorithm, it is possible to infer how the gains determine the weights assigned to each observation included in the estimation (the precise relation will be presented below, in Eqs. (10) and (11)). Namely, the higher the gain, the higher the emphasis given to more recent observations; hence, higher gains imply a smaller memory of past observations in the current estimates provided by the learning algorithm.

Finally, the learning gain can be assumed to be determined as a choice by the agents adopting a given learning algorithm to update their expectations. Approaches following along these lines often require additional behavioral assumptions to discipline such a choice. Examples include: the adaptively rational expectations approach of Brock and Hommes (1997), where agents are assumed to select between a set of predictors according to a fitness measure; and, the "expectation calculation" approach of Evans and Ramey (1998), where agents are assumed to face a calculation decision pondering between the benefits and the costs of extra cognitive efforts to form expectations.

### 2.2. Applied specifications

Three of the main alternatives for the specification of the learning gain are those of a time-decreasing, a time-constant, and a time-varying (not restricted to be decreasing) sequence of values, and their suitability depends on the time-varying nature of the environment. A decreasing-gain was the seminal choice in the learning literature, so as to match the recursive form of the well known Ordinary Least Squares (OLS) estimator from basic econometrics. For the case of linear models with time-invariant parameters, this estimator is known to possess some well desired properties, such as consistency and efficiency, though these properties do not extend to a time-varying context.

Extensive evidence (see, e.g., Stock and Watson, 2003; Cogley and Sargent, 2005) favoring time-varying parameter models of the economy has, nevertheless, challenged this paradigm, and the departure from the parameter constancy assumption has naturally led to the requirement of adjustments to the learning rules as well. The constant gain specification has been in the spotlight of most applied research since Sargent (1999), given its tracking capabilities and its suitability for time-varying environments. More recently, alternative approaches of time-varying gains also have been proposed to deal with relatively stable systems that occasionally undergo structural breaks (see Marcet and Nicolini, 2003; Kostyshyna, 2012; Milani, 2014); one key issue in the introduction of time-varying gains is to find an appropriate statistical foundation for the gain adjustment mechanism, an issue we treat with more detail in Section 5.

Finally, different approaches have been proposed in the recent applied literature to obtain empirical estimates of the learning gain. The main distinctive feature in these approaches refers to the joint estimation of a constant gain with other model's structural parameters; both Bayesian (Milani, 2007, 2008) and classical (Chevillon et al., 2010) estimation approaches

have been successfully applied, though inference is strongly affected by weak identification issues. Particularly, one major challenge to this approach is that as models become more sophisticated and incorporate more realistic behavioral features, such as adaptive learning, the number of parameters determining the dynamics of the model also grows. This, in turn, dangerously undermines the reliability of likelihood-based estimation methods: there are too many degrees of freedom to match the model's implied macro predictions to the data.

The alternative approach, which we also advocate here,[5] is given by calibration procedures attempting to optimize the fit of learning rules in forecasting macroeconomic data and in matching forecasts from surveys. One earlier application of this approach is provided by Orphanides and Williams (2005) in a study of US monetary policy during the 1970s; Branch and Evans (2006) also provide gain calibrations to US data on inflation and output growth, which have been recently extended to other variables by Markiewicz and Pick (2014), and to other countries focusing on inflation by Molnár and Reppa (2010) and Weber (2010) provides a similar application to European data; and Ormeño and Molnár (2015) show the usefulness of information contained in survey data for the estimation of structural models with adaptive learning.

## 3. Adaptive learning context

Macroeconomic models imply relationships between aggregate variables such as inflation, output growth and interest rates, and often involve also past and future expected values of these same variables. In the latter case, one therefore has to specify how expectations are formed, that is, one has to specify a forecasting model (or PLM) for agents. Perhaps one of the greatest virtues of the RE paradigm is the containment of the set of such admissible models by the disciplinary assumption that expectations must be consistent with the implied actual outcomes. The introduction of adaptive learning, in contrast, opens the way for a plethora of (mis)specification possibilities, particularly in multivariate contexts. Which variables should enter agents' PLMs? How does the calibration of the learning mechanism depend on the statistical properties of these variables?

Consider for example a reduced form model given by

$$y_t = \alpha + \beta E_{t-1}^*[y_t] + \delta x_{t-1} + u_t, \tag{1}$$

where $y_t$ is an endogenous variable, $x_t$ is a zero mean exogenous variable, and $u_t$ is an unobservable white noise disturbance. With $\beta < 0$, this is a standard specification of Muth's (1961) partial equilibrium cobweb model of supply and demand, taking $y_t$ as the market clearing price and $x_t$ as a supply shock. With $\beta > 0$, it represents instead a Lucas (1973) aggregate supply model. Multivariate extensions and alternative timing assumptions are possible to fit more complex general equilibrium models. The minimum state variable (MSV) RE equilibrium associated to (1) can be obtained by assuming agents include a constant term and the lagged exogenous variable in their PLM, i.e.,

$$E_{t-1}^{RE}[y_t] = a + b x_{t-1}, \tag{2}$$

where RE dictates that $a = (1 - \beta)^{-1}\alpha$ and $b = (1 - \beta)^{-1}\delta$. Thus, the assumption of RE establishes an unique trajectory of $y_t$ that satisfies (1) for given parameters and a sequence of stochastic shocks $u_t$.

Under adaptive learning, instead, one relaxes the rationality requirement and assumes that agents act like econometricians. There are two important features in the shift from the RE to the adaptive learning approach. First, under adaptive learning agents are not certain about their PLM's parameter values, and hence follow a real-time process to obtain estimates relevant for their expectations formation at the same time that such expectations are determining actual outcomes. Second, bounded rationality features may restrict agents' perceptions about the relevant variables to be included in their PLMs. The first feature is commonly accomplished by assuming agents update their PLM parameter estimates using a recursive LS algorithm of the form

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \gamma_t \mathbf{R}_t^{-1} \mathbf{x}_t (y_t - \mathbf{x}_t' \hat{\boldsymbol{\theta}}_{t-1}), \tag{3}$$

$$\mathbf{R}_t = \mathbf{R}_{t-1} + \gamma_t (\mathbf{x}_t \mathbf{x}_t' - \mathbf{R}_{t-1}), \tag{4}$$

where $\gamma_t$ is a learning gain parameter, $\mathbf{R}_t$ stands for an estimate of regressors matrix of second moments, and $\mathbf{x}_t$ and $\hat{\boldsymbol{\theta}}_t$ collect the variables included in the PLM and their associated parameter estimates, respectively; in our example model, under a PLM of the form of (2), $\mathbf{x}_t = (\,1 \quad x_{t-1}\,)'$ and $\hat{\boldsymbol{\theta}}_t = (\,a_t \quad b_t\,)'$. Substituting this into the reduced form model one finds that the actual law of motion (ALM) of $y_t$ under adaptive learning,

$$y_t = \alpha + \beta a_{t-1} + (\beta b_{t-1} + \delta) x_{t-1} + u_t, \tag{5}$$

now depends on the evolution of the PLM parameter estimates, which are jointly determined by the learning algorithm recursions for given initials, $\hat{\boldsymbol{\theta}}_0$ and $\mathbf{R}_0$, and a sequence of learning gains, $\gamma_t$. For a sufficiently small constant gain, and provided stability conditions on the model parameters and the exogenous variable are satisfied, it is well known that the PLM

---

[5] Our focus on calibration instead of estimation of the learning gain should not debunk the merits of a proper analysis of this issue; these approaches are highly complementary to each other, especially considering that standard numerical estimators still require the definition of a plausible range of values for the parameter to be estimated. Hence, we hope our results to be informative for the estimation of fully-fledged macroeconomic models as well.

parameter estimates will converge to a stationary distribution centered on the RE equilibrium values (Evans and Honkapohja, 2001). Since the statistical properties of this distribution are directly affected by the learning gain, it is important to have some guidance on plausible calibration values for applied purposes.

The convergence of adaptive learning towards the RE equilibrium is nevertheless conditional on the assumption of a correctly specified PLM. Consider for instance what would happen in our example if agents did not realize that $y_t$ depends on the lagged values of $x_t$, assuming instead a naive PLM containing only a constant term. Clearly, such a PLM automatically rules out a RE solution when $\delta \neq 0$. Under learning, however, one can still find support for this PLM misspecification as a restricted perceptions (RP) equilibrium. Substituting the implied ALM,

$$y_t = \alpha + \beta a_{t-1} + \delta x_{t-1} + u_t, \tag{6}$$

into the learning algorithm, (3), one finds that

$$a_t = a_{t-1} + \gamma_t(\alpha + \beta a_{t-1} + \delta x_{t-1} + u_t - a_{t-1}) \tag{7}$$

is a stochastic recursive algorithm governing the evolution of agents' expectations. For sufficiently small learning gains, the mean dynamics for $a_t$, represented by

$$\dot{a} = \lim_{t \to \infty} E(\alpha + (\beta - 1)a_{t-1} + \delta x_{t-1} + u_t) \tag{8}$$

converges to a distribution centered on $a = (1 - \beta)^{-1}\alpha$, thus defining a RP equilibrium.

A dimension of particular interest for the calibration of adaptive learning algorithms therefore relates to the choice of variables to include in agents' information set. Such choice is particularly important in applied works, where it is not necessarily superior to assume that agents are aware of the correct specification for their PLM. In fact, one of the main earlier criticisms to the RE paradigm has been the often unrealistic assumption that agents possess a correct perception of the economy's structure (e.g., Shiller, 1978). For this reason, our calibration analysis will focus on the performance of several PLMs that may be of common use for different macroeconomic models.

## 4. Empirical framework

### 4.1. Model and estimation

Consider an estimation context faced by a real-time agent wishing to obtain inferences about the law of motion of a set of variables of interest, say $\mathbf{y}_t = (y_{1,t}, \ldots, y_{n,t})$. From an economic perspective, these inferences can be thought of as the middle step agents undertake in a process of learning-to-forecast in order to form their expectations. To narrow down our focus, we assume this agent believes that those variables are statistically related to their own lagged values through a linear regression of the form

$$\mathbf{y}_t = \mathbf{x}_t'\boldsymbol{\Theta}_t + \boldsymbol{\epsilon}_t, \tag{9}$$

where $\mathbf{x}_t = (1, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p})'$ is a vector collecting a constant term plus the $np$ lagged values of the endogenous variables, $\boldsymbol{\Theta}_t = (\boldsymbol{\theta}_{0,t}, \boldsymbol{\theta}_{1,t}, \ldots, \boldsymbol{\theta}_{np,t})'$ stands for a $(np+1) \times n$ matrix of (possibly time-varying) coefficients associated to each regressor in $\mathbf{x}_t$, and $\boldsymbol{\epsilon}_t$ denotes a vector of uncorrelated white noise disturbances. For $n > 1$, (9) is commonly referred as an $n$-variables vector autoregression of order $p$, or $n$-VAR($p$) for short. Both coefficients and disturbances are assumed not to be directly observable by the agent, hence a technique for the estimation of $\boldsymbol{\theta}_t$ is required to allow the construction of the inferences of interest.

As it turns out, the LS algorithm of (3) and (4) provides a recursive form of the equation-by-equation[6] estimator obtained by the minimization of the sum of weighted error squares,

$$\hat{\boldsymbol{\theta}}_t = \arg\min \sum_{i=1}^{t} \beta(t, i)(y_i - \mathbf{x}_i'\hat{\boldsymbol{\theta}}_t)^2, \tag{10}$$

where the weights given to each observation are determined by the sequence of gains according to

$$\beta(t, i) = \begin{cases} \dfrac{\gamma_i}{\gamma_t} \displaystyle\prod_{k=i+1}^{t} (1 - \gamma_k) & \text{for } i < t, \\ 1 & \text{for } i = t. \end{cases} \tag{11}$$

Under a constant gain specification, $\beta(t, i) = (1 - \bar{\gamma})^{t-i}$, so that past observations are given geometrically decaying weights, whereas a decreasing gain leads to the famous OLS estimator of basic econometrics (see Berardi and Galimberti, 2013, for

---

[6] For simplicity we abuse notation and let $\hat{\boldsymbol{\theta}}_t$ stand for column $j$ in $\boldsymbol{\Theta}_t$, i.e., the vector of coefficients associated to the $j$th equation in the VAR system; similarly, $y_t$ stands for $y_{j,t}$.

derivations). These properties may provide an explanation for the prominence of the LS algorithm in the adaptive learning literature as the choice to represent agents' mechanism of adaptive learning; here we follow such practice and focus our calibration analysis on the LS case.[7]

### 4.2. Data

We consider the inclusion of three key macroeconomic variables in agents' learning-to-forecast problem: inflation, output growth, and interest rate. Following our previous discussion, we evaluate all possible combinations of these variables in the specification of the learning PLMs; namely, one univariate AR($p$) specification for each variable, three bivariate 2-VAR($p$) specifications, and one trivariate 3-VAR($p$) specification. For further robustness we also vary the VAR lag orders from $p = 1$, ..., 4, hence totaling 16 PLM specifications for each variable.

We use US quarterly data on real GNP/GDP and its price index to obtain measures of output growth and inflation, respectively, and the 3-month Treasury Bill rate as a measure of the interest rate. The sample covers the period from 1947q2 to 2013q4, and is organized in a real-time data format with vintages from 1966q1 to 2014q1; i.e., a total of 193 snapshots of what was known on the above variables by a market participant in real-time.

For the purpose of comparing the learning-based forecasts to those provided by survey respondents, we use data from the Survey of Professional Forecasters (SPF); each quarter, this survey asks professional economists to give their forecasts for several macroeconomic variables and also over different forecasting horizons. Here we use the median of the individual forecasts made for a total of five horizons, namely from $t$ (nowcast) to $t + 4$. The SPF data is available from 1968q4 onwards for inflation and output growth, but only from 1981q3 for the interest rates. To close this gap we follow Milani (2011, p. 385) and derive expectations about 1-period-ahead interest rates according to the expectations theory of the term structure.[8]

### 4.3. Further implementation details

In order to implement the LS recursions one needs to first specify how the initial estimates are obtained. Here we adopt a training sample-based method, where an initial portion of the available data is set aside to the application of the algorithm departing from diffuse initials, i.e., $\hat{\theta}_0 = \mathbf{0}_{np \times 1}$ and $\mathbf{R}_0 = \mathbf{0}_{np \times np}$. Berardi and Galimberti (2017) recently surveyed different methods adopted in the literature and reported evidence in favor of the convergence properties provided by the use of a training sample, in comparison with other more model-dependent initials such as equilibrium-related and estimation-based methods. Also in line with the previous applied literature, we use 75 observations for the training sample, which corresponds to the period from 1947q2 to 1965q4.

Estimation and forecasting are then carried out by vintage as follows: first, recursions for each gain calibration are computed departing from the vintage/gain initials until exhaustion of the vintage sample; then, the $t$, ..., $t + 4$ forecasts for each vintage/gain are computed using the last estimates of the model specification, where $t$ stands for the vintage quarter. We repeat these computations for each vintage of data from 1966q1 to 2012q4, which results in a total of 188 forecasts for each combination of gain calibration, model specification, variable, and forecasting horizon. It remains to be specified how the gain calibrations are determined.

## 5. Gain calibrations

In order to approach empirically the gain calibration issue, we adopt two distinct measures of gain selection. Under the first alternative, we assume gains are determined as a choice by agents aiming to optimize their learning performance. A natural measure to select gain calibrations under this assumption is the accuracy of their associated forecasts. We denote the gains selected according to this criterion as *actual-based* calibrations. As for the second selection measure, we take the learning gain as a primitive parameter of agents' learning-to-forecast behavior. An appropriate measure of fit under this alternative requires the observation of agents' actual expectations. For that purpose we use data of survey forecasts as a proxy, calling this gain selection method a *survey-based* calibration.

Different assumptions are possible regarding the use of real-time data, and about the specification of the range of gain values of interest. We consider two computational approaches to obtain these gain values. In the first, we begin by constructing a grid of admissible values, and then proceed by imposing selection rules that represent our different assumptions on the evolution of the gains. The second approach involves the use of an outer mechanism to adaptively adjust the gain in response to changes in the recent performance of the algorithm.

---

[7] A less frequently adopted alternative is provided by the stochastic gradient (SG) algorithm, which, in spite of being computationally simpler than the LS, raises rather challenging issues in its calibration due to a sensitiveness to the data scales. See Berardi and Galimberti (2014) for a comparison between these algorithms, and Galimberti (2013) for details on the calibration of the SG algorithm.

[8] Further details about the data sources and adjustments are provided in the online appendix.

## 5.1. Grid-based gains

We construct a grid of 100 gain values by setting an upper bound on admissible values so as to ensure the algorithm's stability. Hence, our estimation routine is applied to each model specification with 100 different constant gain values. When it comes to represent agents' learning-to-forecast behavior, nevertheless, an unique gain value is required for each time an iteration on the learning algorithm is performed. Under our two selection criteria, i.e., actual-based and survey-based, we pick the period specific gains by minimizing averages of forecast errors and forecast comparison errors, defined as the difference between the forecasts associated to each gain and the actual[9] and survey observations, respectively. Moreover, since we have multiple forecasting horizons available from the surveys, our gain selections are also based on the average errors over these horizons. Letting $z_{t+h}$ stand for the selection measure (either the actuals, $y_{t+h}$, or the surveys, $s_{t+h}$), the gain is obtained according to

$$\gamma_z = \underset{0 < \gamma \le 0.5}{\arg \min} \left\{ T_S^{-1} \sum_{t \in S} H_t^{-1} \sum_{h \in \mathcal{H}_t} (z_{t+h} - \hat{y}_{t+h}(\gamma))^2 \right\}, \tag{12}$$

where $\hat{y}_{t+h}(\gamma)$ is the forecast of $y_{t+h}$ associated with the model estimates obtained with the gain $\gamma$ and using data in vintage $t$, $h$ is the forecasting horizon (usually $\mathcal{H}_t = \{0, \ldots, 4\}$ and $H_t = 5$, except when particular horizons are missing in the survey data), and $S$ is a selection sample of length $T_S$. The definition of this selection sample depends on whether the gains are assumed to remain fixed throughout the sample, or are allowed to vary over time.

### Fixed gains

Fixed gains are selected by minimizing the average forecast (comparison) errors over the full sample of forecasts that we compute, i.e., 188 forecasts covering the period from 1966q1 to 2012q4. Fig. 1 illustrates this selection for the case of univariate specifications on each variable forecasted, comparing their forecasting performance in terms of mean squared forecast errors (MSFE) and mean squared forecast comparison errors (MSFCE) relative to the survey forecasts. The performance of the fixed gains are also compared to the traditional case of a decreasing gain sequence, which is found to provide inferior accuracy and resemblance to the surveys than the best fixed gains for all variables, though the difference is more pronounced for inflation. Selections based on an alternative evaluation sample, from 1981q1 to 2012q4, are also presented; these suggest some potential misspecification in the use of a fixed gain for inflation, since the optimal gain for this variable depends on the selection sample.

Some "spikes" are observed in the forecasting performance curves of output growth and interest rates. These are caused by the activation of a projection facility (PF), which is a mechanism assumed to be coupled to the learning algorithms such that whenever the estimates leave a bounded region in the parameters space the device is activated in order to contain the escape. Naturally, as the gain increases there is more instability and the estimates get closer to the PF activation threshold[10]. Before reaching that threshold, nevertheless, the forecasts become more "explosive", which deteriorates the performance.

Clearly, under the actual-based calibration, the selection using the full-sample of forecasts violates the restrictions of a "fair" out-of-sample forecasting exercise, since it uses information not available for a real-time learner; e.g., an agent forecasting future inflation rates in real-time from 1990q1 does not know how each of the available gain values will perform over 1990q1, 1990q2, and so on. Branch and Evans (2006) circumvent this issue by defining the calibration within a pre-sample. In contrast, the use of the full-sample of forecast comparison errors under the survey-based gain calibration does not present any conflict, as it would just imply that agents hold immutable beliefs about the system they are forecasting. This alternative has also found some applications in the previous literature (see, e.g., Orphanides and Williams, 2005; Pfajfar and Santoro, 2010).

### Recursive gains

Another possibility is to allow the selection of the gain to be recursive, through a minimization of the average forecast (comparison) errors over a rolling window of forecasts. Here we adopt a window length of 60 observations, the first of which covering the period from 1966q1 to 1980q4.[11] Hence, the recursive gains are obtained through a period-by-period application of (12) with $S_t = \{t - 1, \ldots, t - 60\}$.

Fig. 2 presents the evolution of these gain calibrations for the illustrative case of univariate forecasting model specifications for each variable (the adaptive gains will be detailed below). Clearly, the flexibility provided by the recursive optimization of the gain selection leads to the emergence of interesting patterns; particularly, the gains for inflation and output growth tended to increase over the 1990s, a period characterized by reduced volatility in these variables compared to the preceding decades (during the 1970s and early 80s the US went through the Great Inflation period). The higher gains

---

[9] Following the real-time spirit of learning, the actual-based forecast errors are computed in relation to the first data measures available for each variable.

[10] The projection facility is activated only if the (9)'s companion matrix eigenvalue with highest modulus gets bigger than a critical threshold. Whereas a value of one is the standard requirement for stationarity (see Lütkepohl, 2005, pp. 13–18), we adopt a critical value of 2 so as to allow the algorithm some space to return to the stability region on its own.

[11] Recall we take the first 75 observations (1947q2-1965q4) of our sample for the initialization of the algorithm.
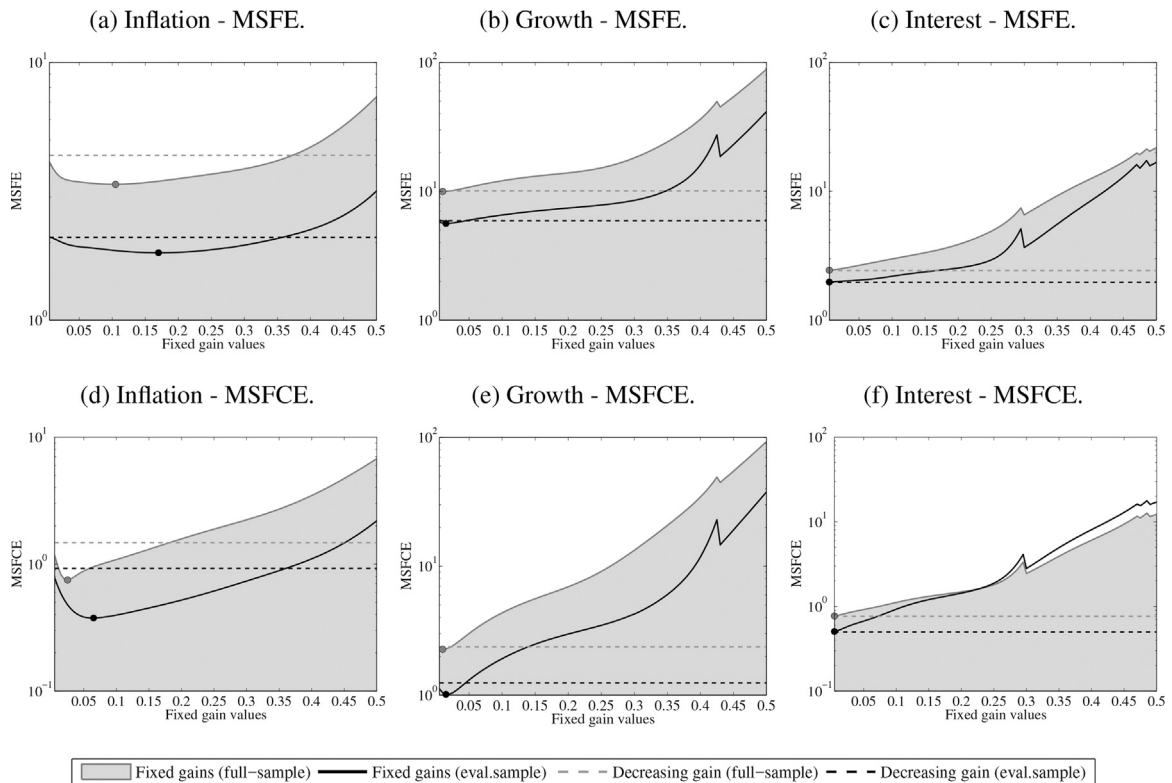
**Fig. 1.** Forecasting performance curves. The forecasts are based on AR(1) univariate specifications for each variable. Performance is measured by the sample average of the squared forecast (comparison) errors averaged over the forecasting horizons. The bullet points indicate the fixed gains that minimize the performance criterion over the corresponding samples (full-sample: 1966q1-2012q4; evaluation sample: 1981q1-2012q4). The major "spikes" in the forecasting performance curves of output growth and interest rates are caused by activations of the projection facility at particular periods (2006q1 for the growth spikes; 1992q1-q2 for the interest rates) for high gain values, as explained in the text. Performance curves for the other specifications are available online in the form of interactive plots.

during the 1990s are consistent with the observation that the economy was undergoing (or had recently undergone) through structural changes that were reflected into the performance of past estimates of the forecasting models.[12]
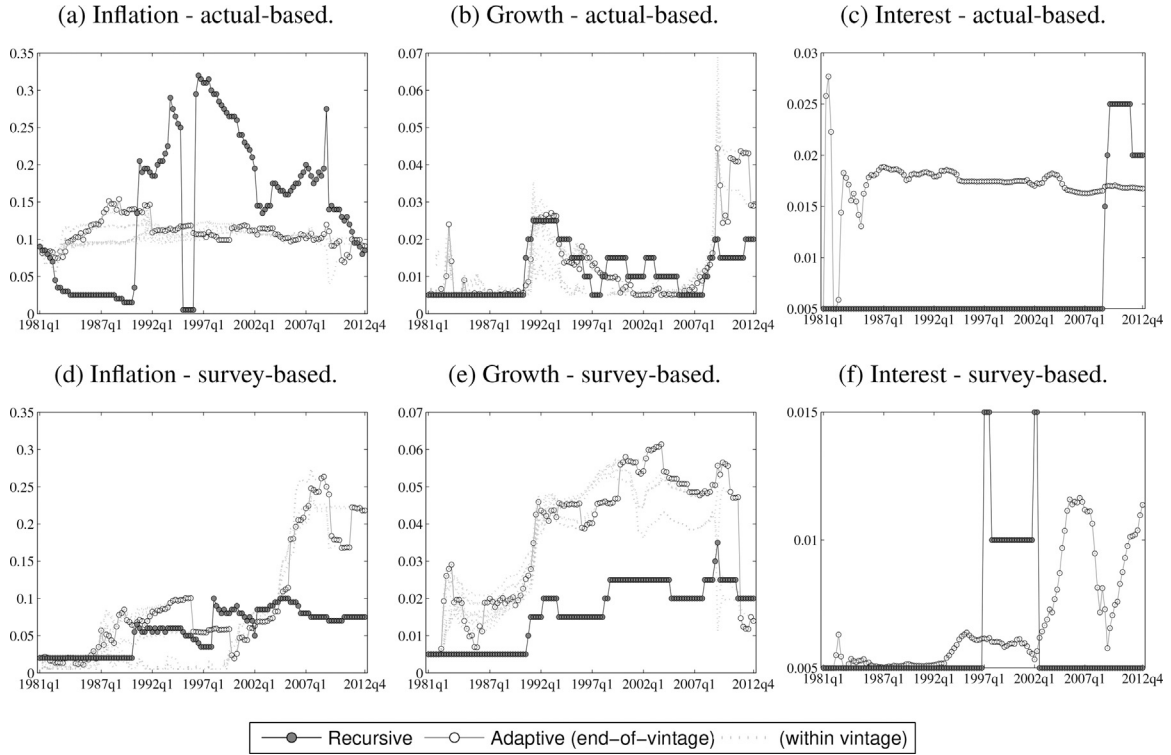
Also notice there are more differences between the actual-based and the survey-based recursive calibrations than observed under the fixed gain assumption. The sequence of gains selected using this recursive approach provides a more realistic approximation to the informational environment faced by an agent learning and forecasting in real-time. Moreover, there is no reason to restrict the gains to be fixed throughout the whole forecasting sample, particularly considering the evidence of misspecification for the case of inflation reported above. Notice, however, that the gain is still kept fixed within each vintage of data used in the calculation of past forecasting performance. E.g., the recursive gain selected to estimate the forecasting model from 1990q1 vintage of data is obtained by minimizing the average forecast (comparison) errors based on the end-of-vintage-sample forecasts of 60 different vintages from 1975q1 to 1989q4. The next approach relaxes this assumption by allowing within vintage time-variation of the gain.

### 5.2. Adaptive gains

An alternative approach to the selection of the learning gains in real-time is to turn the gain calibration itself into an adaptive estimation problem. The idea is to complement the learning algorithm with an outer mechanism to adjust the gain in response to the algorithm's recent performance. Such an automatic tuning of the recursive algorithms was first suggested in Benveniste et al. (1990), and later analyzed by Kushner and Yang (1995) who presented evidence favoring this approach.

---

[12] Considering that higher gains are associated with stronger time-variation in parameter estimates, the increase in the gains during a period of lower volatility may seem puzzling at first. Notice, however, that the relationship between a variable's volatility and the learning gain is not necessarily positive. From a filtering point of view, the optimal gain is tied to the signal-to-noise ratio of the process under estimation, which can be affected by two main factors: first, the degree of time variation in the model's parameters; second, the magnitude of noise affecting the observables. Whereas in the former, greater parameters variation would require higher gains to track the evolution of the parameters, in the latter, higher noise would reduce the quality of the signal and make lower gains optimal. Clearly, both effects can be associated with increased volatility in the variable, but not necessarily with an increase in the gains.

### (a) Inflation - actual-based.

### (b) Growth - actual-based.

### (c) Interest - actual-based.

### (d) Inflation - survey-based.

### (e) Growth - survey-based.

### (f) Interest - survey-based.



—●— Recursive    —○— Adaptive (end–of–vintage)    ⋯⋯ (within vintage)

**Fig. 2.** Evolution of time-varying gain calibrations. The gains are selected in real-time under AR(1) univariate specifications for each variable. The recursive gains are selected on the basis of the forecasting performance (actual or survey-based) of fixed gains over a window of past 60 observations. Adaptive gains are time-varying within each vintage, but with gain values calculated by an outer algorithm instead of a selection based on a grid; the adaptation constants are 0.001, for inflation, 0.0001, for growth, and 0.0005, for interest.

In economics, a recent application has been presented in Kostyshyna (2012), evidencing some quantitative improvements to model hyperinflation episodes based on a PLM including only a constant.

The derivation of the adaptive gain mechanism starts with the specification of an objective function, which in our context is related to the loss functions we associate to the two alternative interpretations given to the learning gains. Hence, the adaptive gain approach is also split in two, with the actual-based gains determined by a mechanism responding to the algorithm's accuracy, while the survey-based gains are driven by the algorithm's resemblance to the survey forecasts. Assuming a squared loss function, the gain is selected by minimizing the expected loss of forecast (comparison) errors

$$J_t = \frac{1}{2}E[(z_t - \mathbf{x}_t'\hat{\theta}_{t-1})^2]. \tag{13}$$

If the true data generating process (DGP) of $z_t$ were known beforehand, one could ideally pick each algorithm's gain so as to optimize the above criterion. In the lack of this information, as it is the case in the learning-to-forecast situation, one alternative is to superimpose an outer adaptive scheme for the purpose of automatic tuning the gain parameter.

The general idea is to use a recursive updating scheme that corrects the gain parameter in the direction opposite to a stochastic approximation of the loss function gradient. Deriving this gradient and plugging in its stochastic approximation we obtain the adaptive gains recursions,[13]

$$\gamma_t = [\gamma_{t-1} + \alpha\mathbf{x}_t'\hat{\boldsymbol{\Psi}}_{t-1}(z_t - \mathbf{x}_t'\hat{\theta}_{t-1})]_{\bar{\gamma}_{\min}}^{\bar{\gamma}_{\max}}, \tag{14}$$

$$\hat{\boldsymbol{\Psi}}_t = (\mathbf{I} - \gamma_t\mathbf{R}_t^{-1}\mathbf{x}_t\mathbf{x}_t')\hat{\boldsymbol{\Psi}}_{t-1}\ldots$$
$$+(\mathbf{I} - \gamma_t\mathbf{R}_t^{-1}\hat{\mathbf{S}}_t)\mathbf{R}_t^{-1}\mathbf{x}_t(y_t - \mathbf{x}_t'\hat{\theta}_{t-1}), \tag{15}$$

$$\hat{\mathbf{S}}_t = (1 - \gamma_t)\hat{\mathbf{S}}_{t-1} + \mathbf{x}_t\mathbf{x}_t' - \mathbf{R}_{t-1}, \tag{16}$$

---

[13] See the online appendix for these derivations. It is important to note that we assume that the derivative of $y_t$ with respect to the gain is null, which neglects the effect of the gain, through the update of expectations, over the determination of $y_t$. Relaxing this assumption goes beyond the scope of this paper.

where $\alpha$ is an adaptation constant, $\hat{\boldsymbol{\Psi}}_t$ stands for an estimate of $\partial\hat{\boldsymbol{\theta}}_t/\partial\gamma$, $\hat{\mathbf{S}}_t$ stands for an estimate of $\partial\mathbf{R}_t/\partial\gamma$, and $[\bullet]_{\bar{\gamma}_{\min}}^{\bar{\gamma}_{\max}}$ is a truncation operator setting $\gamma_t$ to $\bar{\gamma}_{\min}$ if it falls below this value, or to $\bar{\gamma}_{\max}$ if it rises above this value. The remaining components follow from the previous definition of the LS algorithm.

An intuition for this gain adaptation mechanism follows directly from its interpretation as a numerical optimization method, browsing along the error-performance curve in search for an optimal gain. Here $\hat{\boldsymbol{\Psi}}_t$ keeps track of the algorithm's past estimation performance, accumulating its past gradients discounted according to forgetting factors implied by the evolution of the gains. This synthetic measure of past performance is taken as a reference in the gain update Eq. (14): when the latest gradient points towards the same (a different) direction as of $\hat{\boldsymbol{\Psi}}_t$, the adaptive mechanism interprets this as evidence of systematic (contradictory) mistakes; hence the gain is increased (decreased) to intensify (lessen) the algorithm's response to last period error.

Nevertheless, the above interpretation becomes less compelling under the survey-based gains, where $z_t \equiv s_t$. In this case the gain adaptation mechanism has to bear with two distinct estimation objectives: the algorithm's accuracy performance, and its resemblance to the surveys. Following the above interpretation, that means the adaptation mechanism has to browse along two error-performance curves at the same time. Whereas the latest gradient estimate is drawn from the resemblance to surveys objective, the past gradients synthesized in $\hat{\boldsymbol{\Psi}}_t$ refer to the algorithm's accuracy performance. Hence, the gain is increased to intensify the algorithm's response to last period error if its recent performance indicates systematic mistakes on both dimensions, and vice versa.

Computation of the adaptive gains still requires specification of the adaptation constant, $\alpha$. According to the analysis of Kushner and Yang (1995), stability requires this parameter to be small, so as to satisfy $0 < \alpha \leq \bar{\gamma}_{\min}$, together with an appropriate upper bound to the gains. For our purposes we take the extreme gain values in the grids defined above as the bounds for the adaptive gains. Most importantly, Kushner and Yang (1995) also present simulation evidence indicating that the algorithm's performance is not as sensitive to the calibration of $\alpha$ as it is to the learning gain. Therefore, our calibration of these adaptation constants trades off performance optimization in favor of a calibration presenting a sequence of gains not too jumpy, but neither constant.[14]

The resulting sequences of adaptive gains are presented in Fig. 2. Comparison with the previous approach shows that the adaptive gains tend to move in tandem with the recursive gains for inflation and growth, though this comovement is more clear when comparing the recursive gains to the within vintages gain variations of the adaptive approach (dotted gray lines in Fig. 2). The correlation coefficients between the actual-based (survey-based) recursive and the adaptive gains obtained within the last vintage in our sample equal to 0.63 (0.59) and 0.35 (0.66) for inflation and growth, respectively. For interest rates there is no evidence of a relationship between these approaches (correlations of $-0.04$ and $-0.16$ for the actual- and the survey-based gains, respectively), a result we associate with the generally low gains obtained for this variable.

### 5.3. Summary and related literature

Other than the variables forecasted and the selection measures, we obtain gain calibrations over two additional dimensions related to the model specification: the explanatory variables and the number of lags included. Averaged gain calibrations over the first of these dimensions are reported in Table 1, showing non-negligible heterogeneity depending on all remaining dimensions. We draw several interesting observations from these results:

1. The gains selected to forecast inflation are overall higher than those used on output growth and the interest rate, i.e., the more recent observations are more important to forecast inflation than for the other variables. An implication of this result is that inflation's DGP exhibits a stronger degree of time-variation, hence the need for a higher gain to track the model's coefficients.
2. The gains tend to decrease as the number of lags included in the forecasting model specification increases, which may be an indication of model misspecification in the low order VARs; particularly, the omission of relevant lagged observations in the model specification may lead to dynamics resembling to non-stationarity, which, in turn, is captured by the higher learning gains.
3. Overall, the actual-based gains are higher than the survey-based ones, which means that the professional forecasters update their beliefs to incoming data systematically more slowly than would be optimal under the MSFE loss function.
4. The averaged time-varying gains also tend to be higher than the averaged fixed gains, i.e., the time-varying gains tend to increase more often and more strongly than to decrease relative to the fixed full-sample selection. Thus, the flexibility of the time-varying gains provides a useful mechanism to speed up the adjustment of expectation models to abnormal episodes.

Most of these findings are consistent with the gain values adopted in the previous literature, as summarized in Fig. 3. There are two main approaches to the determination of the learning gains in applied works: first, the gain can be jointly

---

[14] We also found that as the number of estimated parameters increased, the adaptive gains turned more unstable for a fixed adaptation constant. To deal with that we set $\alpha = \bar{\alpha}/(np+1)$, where $np+1$ is the number of parameters in the model specification.

**Table 1**
Averaged gain calibrations.

| Calibration | Actual-based calibrations | | | Survey-based calibrations | | |
|---|---|---|---|---|---|---|
| Lag order | Inflation | Growth | Interest | Inflation | Growth | Interest |
| (a) Fixed gains | | | | | | |
| VAR(1) | 10.8 | 1.8 | 0.5 | 1.6 | 1.4 | 0.5 |
| VAR(2) | 2.4 | 1.3 | 0.5 | 0.9 | 1.1 | 0.5 |
| VAR(3) | 2.3 | 0.9 | 0.5 | 0.8 | 0.9 | 0.5 |
| VAR(4) | 3.0 | 1.3 | 0.5 | 1.4 | 1.1 | 0.5 |
| Overall | 4.6 | 1.3 | 0.5 | 1.2 | 1.1 | 0.5 |
| (b) Recursive gains | | | | | | |
| VAR(1) | 12.8 | 1.7 | 0.8 | 4.0 | 2.5 | 0.7 |
| VAR(2) | 5.0 | 1.3 | 1.2 | 2.0 | 1.7 | 1.0 |
| VAR(3) | 3.1 | 1.4 | 1.6 | 1.0 | 1.7 | 1.0 |
| VAR(4) | 2.9 | 1.5 | 1.0 | 1.3 | 1.7 | 0.8 |
| Overall | 6.0 | 1.5 | 1.1 | 2.1 | 1.9 | 0.9 |
| (l) Adaptive gains | | | | | | |
| VAR(1) | 7.6 | 1.2 | 1.4 | 4.4 | 2.3 | 0.8 |
| VAR(2) | 3.7 | 1.3 | 0.8 | 3.8 | 1.7 | 1.3 |
| VAR(3) | 5.1 | 1.1 | 1.7 | 5.1 | 1.3 | 0.8 |
| VAR(4) | 3.4 | 1.1 | 0.9 | 4.8 | 1.4 | 1.8 |
| Overall | 4.9 | 1.2 | 1.2 | 4.5 | 1.7 | 1.1 |

The presented gain values are scaled by 100, e.g., 10.8 is equivalent to $\gamma = 0.108$. Averages are based on fixed and averaged time-varying gains selected under different model specifications (univariate, bivariate, and trivariate).
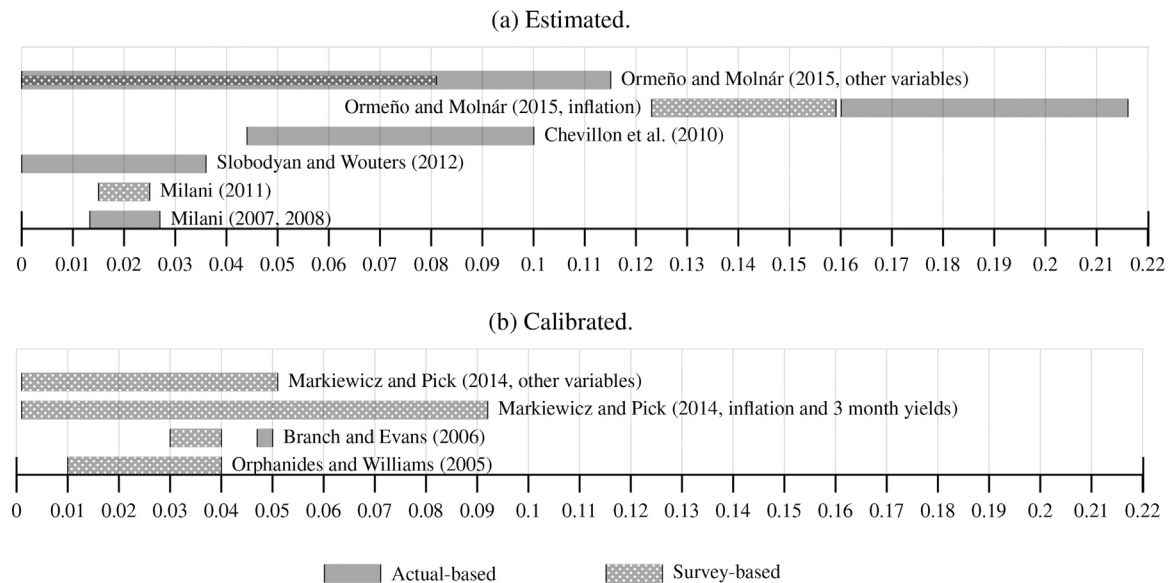


Fig. 3. Gain values in the literature. All gains are based on US data, but for diverse samples and model specifications. The intervals reported for estimated gains refer to 90% credible (Bayesian estimates) and confidence (GMM estimates) intervals, except for Milani (2007, 2008, 2011), who reports 95% credible intervals, and Ormeño and Molnár (2015), who report median and standard deviations of the estimates (intervals are ±2 st.devs. around medians). Most studies adopt a common gain for the different forward variables requiring expectations; those that estimate/calibrate different gains for each variable are indicated.

estimated with other parameters within a structural model; second, the gain can be calibrated according to the forecasting performance. Both instances allow the use of actual observations or survey forecasts as a reference, the latter often yielding lower gain values. Moreover, the use of different gains for each variable has been increasingly popular.[15]

Given all this heterogeneity, an important question is which of these calibrations should be favored in applied research. We now turn to this issue.

---

[15] Calibrations for other countries also present a great diversity of gain values (all on inflation): for European economies, Weber (2010) documents actual-based gains between [0.067, 0.3], experts-survey-based gains between [0.013, 0.178], and households-survey-based gains between [0.001, 0.093]; Molnár and Reppa (2010) find survey-based gains between [0, 0.08], for developed economies, and between [0.04, 0.26] for Eastern European and Latin American economies.

## 6. Forecasting performance

We compare the performance of the different gain calibration assumptions under two evaluation criteria defined according to the measure used for the gain selection. Namely, we evaluate the performance of the actual-based calibrations in terms of their forecasting accuracy, measured by the MSFEs, whereas for the survey-based calibrations we focus on the forecasts resemblance to the surveys, measured by the MSFCEs. The motivation for these criteria naturally arise from the purpose of each calibration. In the first we take the point of view of an economic agent, who has to build forecasts of variables relevant for economic decisions; what matters for this agent is the accuracy of such forecasts. In the second criterion we assume the point of view of the researcher, whose interest, in contrast, is in uncovering which mechanism better represents the learning-to-forecast behavior of the economic agents being modeled.

There is some overlap between these evaluation criteria and the measures we use for the gain selections, particularly for the case of the fixed gains. To circumvent this issue we attempt to reduce the informational advantage provided to the fixed gain calibration, relative to the backward-looking time-varying calibrations, by restricting our evaluation sample. Namely, our evaluation covers the period from 1981q1 to 2012q4, which also aligns with the forecasts obtained under the recursive calibration after the initial window of 60 observations (it also corresponds to the samples presented in Fig. 1). Hence, because the calibrations are based on the gains performance in a past sample, they are not necessarily the optimal gains in the evaluation sample. Under this design, our evaluation can uncover evidence of model misspecification and time-varying features in the underlying DGPs.

### 6.1. Overview

Fig. 4 compares the forecasting performance of the gain calibrations over the several dimensions we are considering. Starting with the evidence in forecasting inflation, panel (a) shows that the univariate model often outperforms the other specifications including additional variables; the optimal lag order also depends on the model specification. Finally, the fixed gain values reported in the middle panel confirm our earlier observation that the selections are mainly determined by the lag orders and the selection measure.

The results on output growth, panel (b) in Fig. 4, reveal a slightly different pattern regarding the optimal lag order, overall favoring the use of only one lag. The univariate model specification is again favored, particularly in terms of resemblance to the surveys; in terms of forecasting accuracy, the trivariate model specification with only one lag is the best, but the performance of this model deteriorates substantially as the lag order increases.

Perhaps more surprising are the results on forecasting the interest rate, presented in panel (c) of Fig. 4, where learning seems to be less relevant: the selected gain calibrations equal the smallest value available in our grid, suggesting little advantages in the use of recursively estimated models to forecast this variable that is more directly affected by central bank policies.

### 6.2. Comparative exercises

Our results above provide clear evidence that the relative performance of the gain calibrations is sensitive to dimensions involving the PLM model specification. In order to obtain a more general picture, we now go beyond those dimensions and synthesize our comparisons in the form of *hit rate* measures. The hit rate stands for the frequency by which the forecasts associated to a given calibration is found to outperform those associated to a competitor with respect to one of our evaluation criteria. Here we conduct paired comparisons between 80 sequences of forecasts associated to each gain calibration, covering combinations of 4 model specifications per variable, 4 lag orders per specification, and 5 forecasting horizons per combination of the former two dimensions. The results are presented in Table 2.
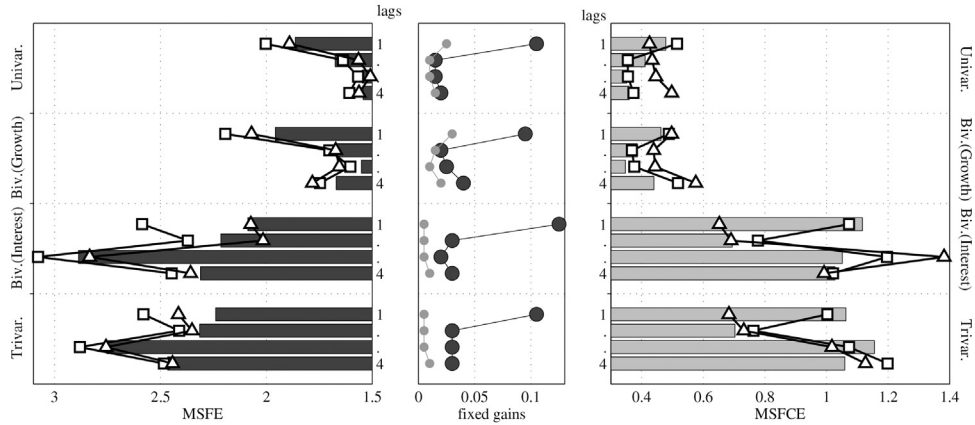
The evidence is quite mixed, depending on the variable forecasted and the evaluation criterion for most of the comparisons. Whereas the actual-based fixed gains tend to outperform the recursive ones in forecasting inflation and output growth, the comparison is leveled against the adaptive gains. To forecast interest rates, the recursive gains seem to be the best choice. In terms of resemblance to the surveys, there is little support in favor of the adaptive gains. The comparison between the survey-based fixed gains and the recursive ones is balanced on inflation and output growth, but dominated by the former in forecasting interest rates.

To further substantiate our comparative analysis, we also make use of tests common to the literature on forecast evaluation. Namely, we adopt both the Diebold and Mariano (1995) (DM) test for equal (unconditional) predictive ability, and its more recently developed conditional counterpart test of Giacomini and White (2006) (GW).[16] Other than for robustness purposes, our choice for these two tests can also be well justified: while the first stands as a classical test, whose properties have been long studied in the literature, the second clearly represents a more appropriate test for our purposes of comparison of different estimation methods.[17]
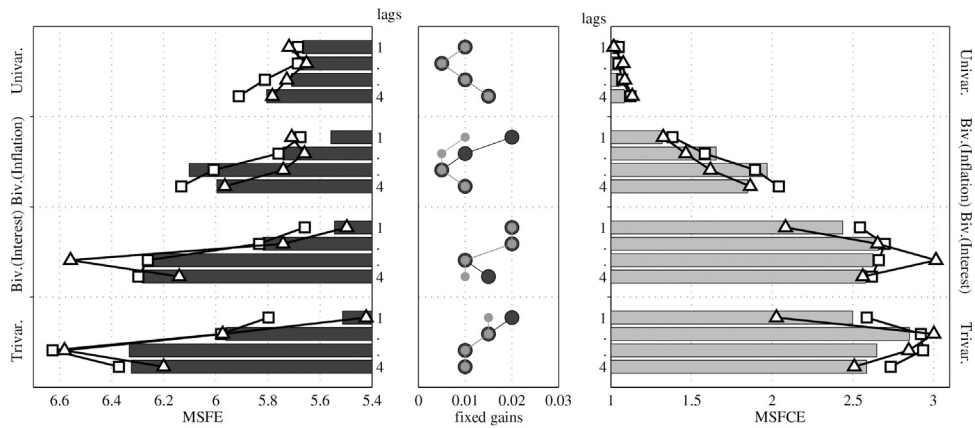
---

[16] We summarize the calculations involved in each of these tests in the online appendix.

[17] Although the GW test is not suitable for the comparison of recursively estimated model-based forecasts, our focus on constant gain specifications attach geometrically decaying weights to past observations, hence approximating a rolling window estimation scheme.
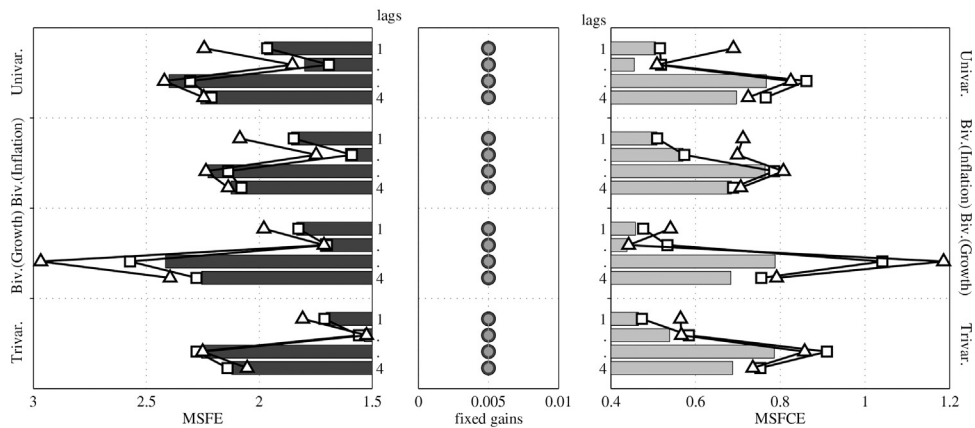
**Fig. 4.** Forecasting performance of gain calibrations. The fixed gains are based on the full-sample, 1966q1-2012q4, whereas the recursive gains are selected over a moving window of past 60 observations. The adaptive gains are based on an outer algorithm with adaptation constants adjusted according to the variable and model specification (see footnote [14]). The reported performance measures are then calculated over the evaluation sample, 1981q1-2012q4.

**Table 2**
Hit rates on comparisons between gain calibrations.

| Statistic | Inflation | | Growth | | Interest | |
|---|---|---|---|---|---|---|
| | A wins | B wins | A wins | B wins | A wins | B wins |
| *Actual-based calibrations – MSFE criterion* | | | | | | |
| *Fixed (A) × Recursive (B)* | | | | | | |
| Hit-rates | 85.0% | 15.0% | 70.0% | 30.0% | 38.8% | 61.2% |
| DM 10% signif. | (15.0%) | (3.8%) | (2.5%) | (1.3%) | (0.0%) | (13.8%) |
| GW 10% signif. | [2.5%] | [1.3%] | [1.3%] | [1.3%] | [0.0%] | [8.8%] |
| *Fixed (A) × Adaptive (B)* | | | | | | |
| Hit-rates | 58.8% | 41.2% | 42.5% | 57.5% | 71.2% | 28.8% |
| DM 10% signif. | (10.0%) | (6.3%) | (5.0%) | (3.8%) | (5.0%) | (1.3%) |
| GW 10% signif. | [7.5%] | [7.5%] | [0.0%] | [3.8%] | [3.8%] | [2.5%] |
| *Recursive (A) × Adaptive (B)* | | | | | | |
| Hit-rates | 26.2% | 73.8% | 26.2% | 73.8% | 81.2% | 18.8% |
| DM 10% signif. | (1.3%) | (10.0%) | (2.5%) | (3.8%) | (10.0%) | (0.0%) |
| GW 10% signif. | [2.5%] | [8.8%] | [2.5%] | [13.8%] | [2.5%] | [0.0%] |
| *Survey-based calibrations – MSFCE criterion* | | | | | | |
| *Fixed (A) × Recursive (B)* | | | | | | |
| Hit-rates | 56.2% | 43.8% | 48.8% | 51.2% | 68.6% | 31.4% |
| DM 10% signif. | (25.0%) | (13.8%) | (0.0%) | (6.3%) | (0.0%) | (4.3%) |
| GW 10% signif. | [6.3%] | [10.0%] | [5.0%] | [2.5%] | [0.0%] | [0.0%] |
| *Fixed (A) × Adaptive (B)* | | | | | | |
| Hit-rates | 78.8% | 21.2% | 65.0% | 35.0% | 67.5% | 32.5% |
| DM 10% signif. | (36.3%) | (11.3%) | (20.0%) | (3.8%) | (33.8%) | (6.3%) |
| GW 10% signif. | [18.8%] | [6.3%] | [10.0%] | [1.3%] | [1.3%] | [3.8%] |
| *Recursive (A) × Adaptive (B)* | | | | | | |
| Hit-rates | 78.8% | 21.2% | 75.0% | 25.0% | 62.5% | 37.5% |
| DM 10% signif. | (33.8%) | (10.0%) | (42.5%) | (1.3%) | (25.0%) | (7.5%) |
| GW 10% signif. | [17.5%] | [2.5%] | [32.5%] | [12.5%] | [1.3%] | [1.3%] |

The hit rates sum up the paired comparisons of forecasting performance between the gain calibrations over the model specifications (4 per variable), the (V)AR lag orders (4 per specification), and the forecasting horizons (5 per spec./lag combination). The DM and GW hit rates refer to the frequency with which the hypothesis of equal predictive ability between the forecasts being compared was rejected at a 10% level of significance.

The results on these statistical tests are again synthesized in hit rate measures, and are presented in Table 2. For most cases the tests for equal predictive ability suggest there is little difference between the forecasts associated to the alternative gain calibrations; the frequency of statistically significant rejections of the null of equal predictive ability observed with the GW tests, for example, average out to about only 5.6% of the comparisons over all variables and evaluation criteria.[18]

### 6.3. Discussion

The results presented in this section provided mixed evidence to support the choice of a gain calibration. There was a great degree of heterogeneity on the forecasting performance of the gain calibrations, particularly with respect to the variable forecasted and the model specification. Importantly, most of the comparisons were not statistically conclusive, meaning that the predictive ability of the different gain calibrations was hard to distinguish.

One implication of these results is that changes in the learning gains, within the reasonable limits we have established, have little effects over the forecasting performance of recursively estimated models involving macroeconomic variables. In contrast, our analysis highlights a greater impact of the model specification on such forecasting performance, both regarding the choice of variables and of number of lags to include in the forecasting models.

To corroborate this idea, in Fig. 5 we draw uncertainty bands around the averaged forecasts obtained with the different calibrations and model specifications (see the Figure's footnote for details about their construction). Clearly, the variability of forecasts over gain calibrations is in general smaller than that observed across model specifications. The fact that the gain uncertainty bands are thin, compared for instance to the variability in the forecasting targets, helps explain the lack of power presented by the statistical tests in distinguishing between the predictive ability associated to the gain calibrations.

Another interesting question relates to the relevance of the gain calibrations for the performance of learning during transitional periods. To assess these effects, in Fig. 6 we split our evaluation sample in three sub-samples so as to cover three distinct periods in the evolution of the US economy: the final years of the Great Inflation, the Great Moderation period, and the recent financial crisis, respectively.

One first interesting observation from the results in Fig. 6 relates to the forecasting of inflation during the period surrounding the recent financial crisis: whereas the adaptive gains provided the best response to the instabilities of this period in

---

[18] An alternative forecast evaluation exercise is to compare the different learning forecasts to a benchmark random walk model, which in our context is equivalent to so-called static or naive expectations adopted in the early literature. Hit rates for this exercise are reported in the online appendix and provide sharper evidence in favor of the adaptive learning approach for forecasting inflation and output growth, but not for interest rates.
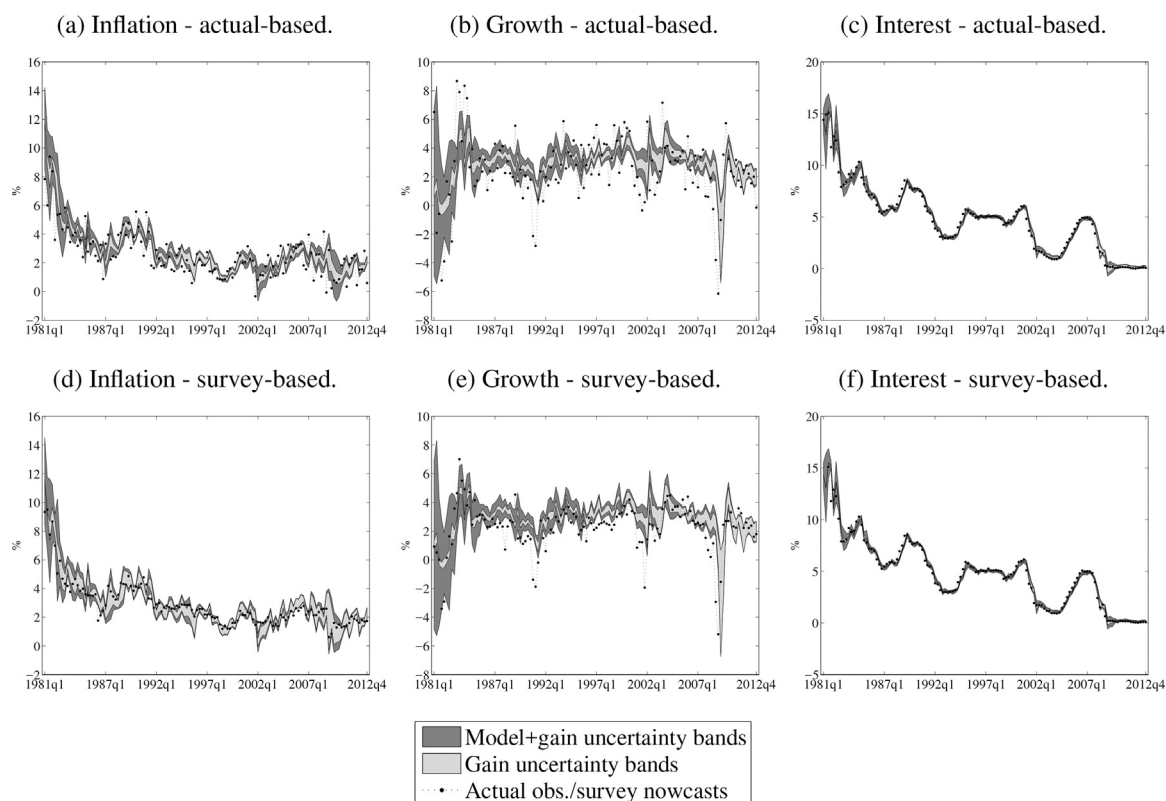
**Fig. 5.** Evolution of nowcasts variability between models and calibrations. The uncertainty bands are ±2 standard deviations around each period's nowcast averaged over the model specifications and the gain calibrations. The uncertainty due only to the gain calibrations are obtained by averaging the standard deviations between the calibrations over the model specifications. The actual observations refer to the first measures observed in real-time.
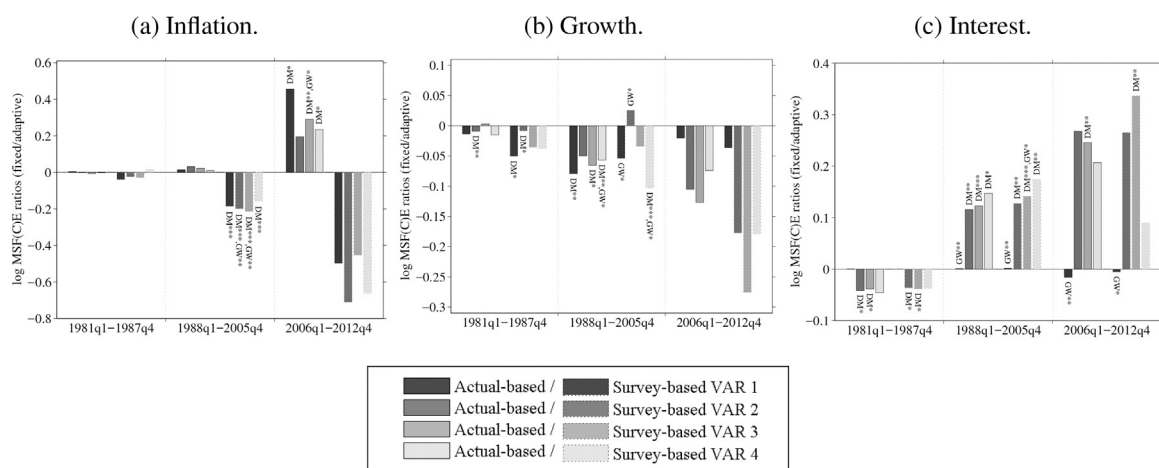


**Fig. 6.** Comparison of nowcasting performance over sub-samples. The nowcasts are based on univariate specifications for each variable. Actual-based (survey-based) calibrations are compared in terms of logs of MSF(C)E ratios between fixed and adaptive gains: positive values indicate the adaptive gains outperformed the fixed gain, and vice versa. Statistical rejections of the hypothesis of equal predictive ability using both the DM and GW tests are indicated by one, two, or three stars, for three corresponding significance levels: 1%, 5%, and 10%.

terms of forecasting accuracy, the fixed gains provided better approximations to the survey forecasts; these results indicate a restrained response from the professional forecasters in the adjustment of their learning process about inflation, suggesting a potential association between the degree of bounded rationality and the degree of uncertainty in the macroeconomy.

For growth, the results in Fig. 6 are rather favorable to the use of a fixed gain over the three sub-samples, and both in terms of forecasting accuracy and resemblance to the surveys; these results, which are also less pronounced than for the other variables, indicate that the process of learning-to-forecast growth is less sensitive to the gain values.
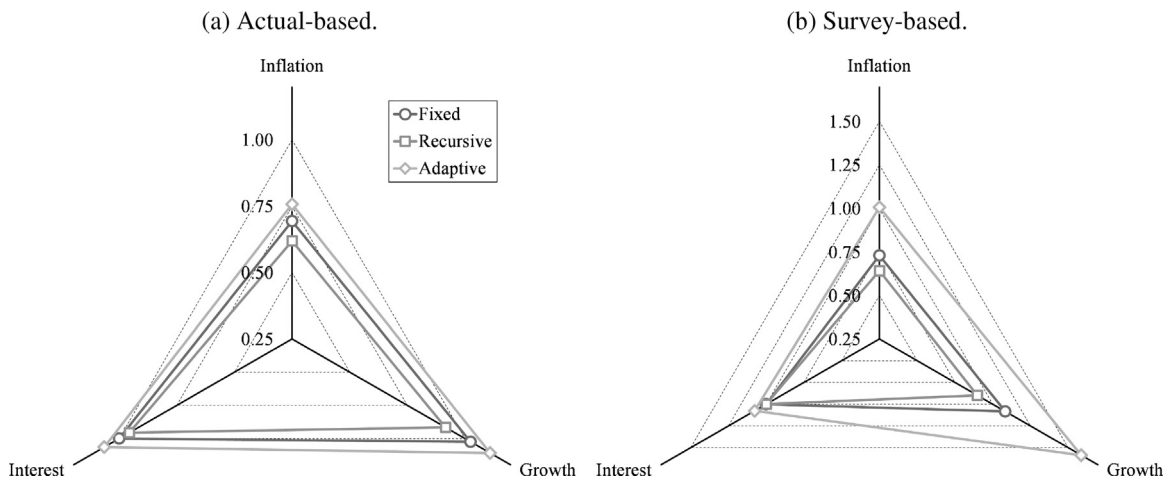
**Fig. 7.** Averaged gain values using revised data relative to real-time calibrations. The plotted values represent ratios between the averaged gain calibrations obtained using revised data and those obtained in real-time. Values below 1 indicate cases where the calibrations for revised data are, on average, smaller than the real-time ones, and vice versa. The fixed gains are averaged over the model specifications and the lag orders, whereas the time-varying specifications are also averaged over the time dimension.

Finally, for interest the adaptive gains have an advantage on both criteria over the last two sub-samples. Thus, in contrast to the learning process for inflation forecasting, market participants seem to be more attentive to the actions of the monetary policy authorities, adjusting their learning for interest rate forecasting more promptly according to the state of the economy; and they are right to do so, as indicated by the results favoring the time-varying gains in terms of forecasting accuracy too.

## 7. Further calibration dimensions

### 7.1. Using revised data

Although the availability and use of real-time data has been increasingly widespread in the applied literature, the traditional approach to the calibration or estimation of macroeconomic models involves the use of only one vintage of data. Whereas we believe the focus on revised data restricts the analysis of learning, by violating the informational restrictions faced by agents when forming their expectations in real-time, it is important to assess its effects over the gain calibrations we have hitherto obtained using real-time data.

Fig. 7 compares the averaged gain values obtained using revised data, from the 2014q1 vintage, to those obtained in real-time.[19] The effects of using revised data are similar across the gain calibrations, depending mainly on the variable forecasted; inflation appears to be the most affected variable, with revised calibrations about 25% smaller than those obtained in real-time; the survey-based adaptive gains also showed some sensitivity to the timing of the data, particularly on growth, where the average gain was more than 50% higher than that obtained in real-time.

It is important to mention that the use of survey forecasts jointly with revised data to calibrate the gains represents a critical violation to the timing of information implied by a realistic learning process. Namely, the forecasts based on learning models estimated with revised data incorporate information that was not available at the time the survey panelists submit their forecasts. Thus, we recommend that the survey-based calibration should be avoided in applications involving revised data.

### 7.2. Alternative variables definitions

Most macroeconomic models are based on relationships for the determination of the conceptual measures of aggregate economic activity that we have explored so far. For applied purposes, nevertheless, there is no unanimous definition of how these concepts should be measured, hence the importance of evaluating how alternative data definitions may affect the calibrations of the learning gains.

In Table 3 we report fixed gain calibrations obtained with alternative data measures that are also commonly employed in the literature, namely, CPI inflation, the unemployment rate, and the 10-year Treasury Bond rate. These are based on univariate specifications of the forecasting model for each variable, hence they are not directly comparable to the averages presented in Table 1, but see the middle panels of Fig. 4. Overall, we observe that only the calibrations for inflation are

---

[19] The averaged gain values calibrated to revised data are presented in the online appendix.

**Table 3**
Fixed gain calibrations for alternative variables.

| Lag order | Actual-based calibrations | | | Survey-based calibrations | | |
|---|---|---|---|---|---|---|
| | CPI inflation | Unemployment rate | 10-year T-bond | CPI inflation | Unemployment rate | 10-year T-bond |
| AR(1) | 3.0 | 0.5 | 0.5 | 4.0 | 0.5 | 0.5 |
| AR(2) | 2.5 | 1.0 | 0.5 | 4.0 | 1.5 | 0.5 |
| AR(3) | 1.5 | 1.0 | 0.5 | 0.5 | 1.0 | 0.5 |
| AR(4) | 3.0 | 1.0 | 0.5 | 2.5 | 1.0 | 0.5 |
| Average | 2.5 | 0.875 | 0.5 | 2.75 | 1.0 | 0.5 |

The presented gain values are based on univariate specifications for each variable, and are scaled by 100.
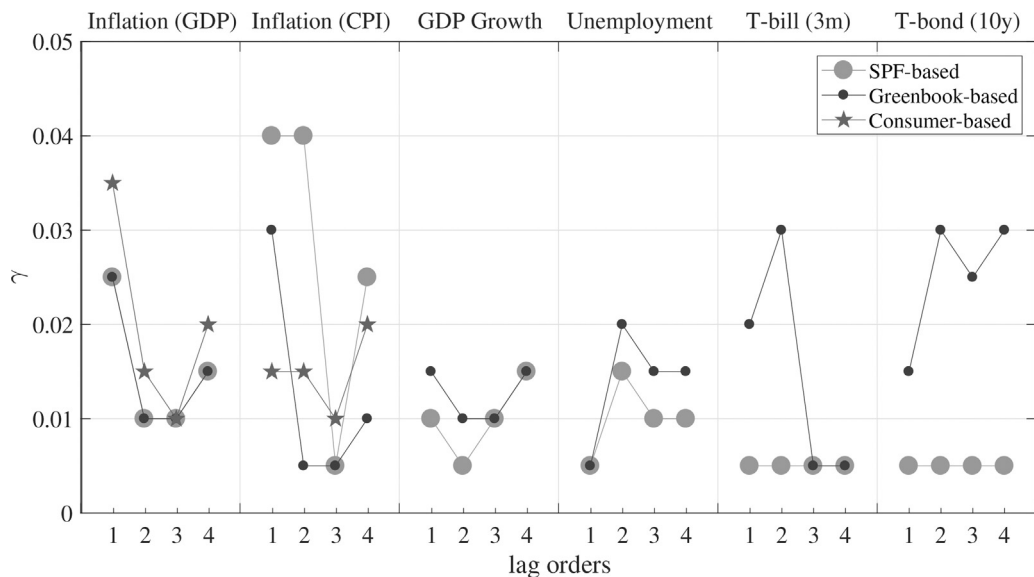


**Fig. 8.** Policymakers and consumers fixed gain calibrations. The presented gain values are based on univariate specifications for each variable and are selected over different sample periods, depending on the availability of the corresponding survey data.

affected by the use of the CPI instead of the GNP/GDP price index; whereas the actual-based gains for this variable decrease, the survey-based gains increase, attenuating the differences between these two calibration approaches.

### 7.3. Matching policymakers beliefs

Another dimension of interest for researchers applying adaptive learning to policy issues relates to calibrations aimed to represent policymakers beliefs (see, e.g., Carboni and Ellison, 2009). For that purpose we re-evaluate our survey-based fixed gain calibrations using Greenbook forecasts. The Greenbooks are reports produced by the research staff of the Federal Reserve Board of Governors before each meeting of the Federal Open Market Committee. They contain projections and assumptions about the future evolution of many economic variables, including those we have analyzed here.

Fig. 8 presents the resulting fixed gain calibrations matching the Greenbook forecasts. For comparative purposes, we also present our earlier survey-based calibrations using the SPF. Interestingly, there are only small deviations from the calibrations obtained to represent the professional forecasters. Except for inflation, the Greenbook-based gains tend to be slightly higher than those based on the SPF; this result suggests that the policymakers adjust their beliefs more quickly than the professionals to new observations on output-related and interest-related variables, but more slowly on inflation.

### 7.4. Matching consumers beliefs

Finally, adaptive learning is often introduced in models where the relevant expectations are associated to households intertemporal decisions. Whereas the professional forecasts have been used in most of the previous literature as a proxy for such expectations, it is important to assess whether learning calibrations may differ when targeting the behavior of consumers. To address this issue we compute the fixed gain calibrations using the Michigan Surveys of Consumers, conducted by the Survey Research Center at the University of Michigan. Each month since January 1978 this survey has been collecting data on different aspects of consumer attitudes and expectations. Here we use the quarterly mean responses on the question

about expected price changes during the next year.[20] In order to accommodate the annual forecasting horizon we obtain the implied adaptive learning forecasts by compounding the first four model-based quarterly forecasts for each period.

The resulting calibrations are presented in Fig. 8 together with the professionals and policymakers calibrations for comparative purposes. The consumers calibrations are obtained for both GDP deflator and CPI inflation forecasts. Interestingly, the consumers gains are in the same range as those obtained for the professionals and policymakers, varying between 0.01 to 0.035. These results indicate that the SPF provides a good approximation to consumers and policymakers beliefs about inflation for the purpose of adaptive learning calibration.

## 8. Concluding remarks

In this paper we have studied empirically the issue of how to calibrate adaptive learning mechanisms to mimic a real-time process of expectations formation in macroeconomic models. We have produced and documented renewed numerical calibrations of the learning gains for applications with US data on inflation, output growth, unemployment and interest rates. We have considered both actuals and survey forecasts data on these variables as reference for our calibrations. We have also explored other dimensions of applied interest, including several combinations of the above variables in the model specifications, and different assumptions regarding the determination of the learning gains.

We have found evidence of a great degree of heterogeneity in the gain calibrations, depending mainly on the variable forecasted and the model specification. The measure used as reference for the selection of the gains also mattered: survey-based calibrations were found to yield lower gain values than those based on actuals. In terms of forecasting performance, the effects of different assumptions regarding the time-variation of the learning gains were found to be of limited statistical relevance. In contrast, the model specification is found to play a greater role in the determination the learning models predictive ability.

Our main conclusion from these results is that an appropriate approach for the determination of the learning gains should be defined in accordance with each model's assumptions: whereas actual-based calibrations should be favored in models aiming to represent rationally optimizing agents, the survey-based calibration should be preferred on bounded rationality grounds; similarly, models allowing for abrupt structural breaks in the evolution of the economic system should consider the possibility of time-varying gains. We hope to have provided some guidance on the design of such mechanisms too.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jebo.2017.10.004.

## References

Barucci, E., 1999. Heterogeneous beliefs and learning in forward looking economic models. J. Evol. Econ. 9 (4), 453–464.
Benveniste, A., Metivier, M., Priouret, P., 1990. Adaptive Algorithms and Stochastic Approximations. Springer-Verlag.
Berardi, M., Galimberti, J.K., 2013. A note on exact correspondences between adaptive learning algorithms and the kalman filter. Econ. Lett. 118 (1), 139–142.
Berardi, M., Galimberti, J.K., 2014. A note on the representative adaptive learning algorithm. Econ. Lett. 124 (1), 104–107.
Berardi, M., Galimberti, J.K., 2017. On the initialization of adaptive learning in macroeconomic models. J. Econ. Dyn. Control 78, 26–53.
Branch, W.A., Evans, G.W., 2006. A simple recursive forecasting model. Econ. Lett. 91 (2), 158–166.
Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. Econometrica 65 (5), 1059–1095.
Bullard, J., Eusepi, S., 2005. Did the great inflation occur despite policymaker commitment to a Taylor rule? Rev. Econ. Dyn. 8 (2), 324–359.
Carboni, G., Ellison, M., 2009. The great inflation and the greenbook. J. Monet. Econ. 56 (6), 831–841.
Carceles-Poveda, E., Giannitsarou, C., 2007. Adaptive learning in practice. J. Econ. Dyn. Control 31 (8), 2659–2697.
Chevillon, G., Massmann, M., Mavroeidis, S., 2010. Inference in models with adaptive learning. J. Monet. Econ. 57 (3), 341–351.
Cho, I.-K., Williams, N., Sargent, T.J., 2002. Escaping Nash inflation. Rev. Econ. Stud. 69 (1), 1–40.
Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. J. Polit. Econ. 113 (1), 1–45.
Cogley, T., Sargent, T.J., 2005. Drifts and volatilities: monetary policies and outcomes in the post WWII US. Rev. Econ. Dyn. 8 (2), 262–302.
Curtin, R.T., 2004. Psychology and macroeconomics: fifty years of the surveys of consumers. In: House, J.S., Juster, F.T., Kahn, R.L., Schuman, H. (Eds.), A Telescope on Society: Survey Research and Social Science at the University of Michigan and Beyond. University of Michigan Press.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13 (3), 253–263.
Eusepi, S., Preston, B., 2011, October. Expectations, learning, and business cycle fluctuations. Am. Econ. Rev. 101 (6), 2844–2872.
Evans, G.W., Honkapohja, S., 2001. Learning and expectations in macroeconomics. In: Frontiers of Economic Research. Princeton University Press, Princeton, NJ.
Evans, G.W., Honkapohja, S., Williams, N., 2010. Generalized stochastic gradient learning. Int. Econ. Rev. 51 (1), 237–262.
Evans, G.W., Ramey, G., 1998. Calculation, adaptation and rational expectations. Macroecon. Dyn. 2 (02), 156–182.
Galimberti, J.K., 2013. Adaptive Learning for Applied Macroeconomics (Ph. D. thesis). The University of Manchester.
Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74 (6), 1545–1578.
Honkapohja, S., Mitra, K., 2006. Learning stability in economies with heterogeneous agents. Rev. Econ. Dyn. 9 (2), 284–309.
Huang, K.X., Liu, Z., Zha, T., 2009. Learning, adaptive expectations and technology shocks. Econ. J. 119 (536), 377–405.

---

[20] The actual survey question about inflation reads as follows: "By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?". The survey also contains questions about interest rates, unemployment and economic conditions expectations, but these are not framed in an open-ended format that can be directly associated with the variables from our analysis (see Curtin, 2004, for more details about the survey).

Kostyshyna, O., 2012. Application of an adaptive step-size algorithm in models of hyperinflation. Macroecon. Dyn. 16, 355–375.
Kushner, H., Yang, J., 1995. Analysis of adaptive step-size SA algorithms for parameter tracking. IEEE Trans. Autom. Control 40 (8), 1403–1410.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer.
Lucas, R.E.J., 1973, June. Some international evidence on output-inflation tradeoffs. Am. Econ. Rev. 63 (3), 326–334.
Marcet, A., Nicolini, J.P., 2003. Recurrent hyperinflations and learning. Am. Econ. Rev. 93 (5), 1476–1498.
Markiewicz, A., Pick, A., 2014. Adaptive learning and survey data. J. Econ. Behav. Organ. 107 (Part B), 685–707 (Empirical Behavioral Finance).
Milani, F., 2007. Expectations, learning and macroeconomic persistence. J. Monet. Econ. 54 (7), 2065–2082.
Milani, F., 2008. Learning, monetary policy rules, and macroeconomic stability. J. Econ. Dyn. Control 32 (10), 3148–3165.
Milani, F., 2011. Expectation shocks and learning as drivers of the business cycle. Econ. J. 121 (552), 379–401.
Milani, F., 2014. Learning and time-varying macroeconomic volatility. J. Econ. Dyn. Control 47, 94–114.
Molnár, K., Reppa, Z., 2010, October. Economic Stability and the Responsiveness of Inflation Expectations. Mimeo.
Muth, J.F., 1961. Rational expectations and the theory of price movements. Econometrica 29 (3), 315–335.
Ormeño, A., Molnár, K., 2015, June. Using survey data of inflation expectations in the estimation of learning and rational expectations models. J. Money Credit Bank. 47 (4), 673–699.
Orphanides, A., Williams, J.C., 2005, November. The decline of activist stabilization policy: natural rate misperceptions, learning, and expectations. J. Econ. Dyn. Control 29 (11), 1927–1950.
Pfajfar, D., Santoro, E., 2010. Heterogeneity, learning and information stickiness in inflation expectations. J. Econ. Behav. Organ. 75 (3), 426–444.
Primiceri, G.E., 2006. Why inflation rose and fell: policy-makers' beliefs and U.S. postwar stabilization policy. Q. J. Econ. 121 (3), 867–901.
Sargent, T., Williams, N., 2005. Impacts of priors on convergence and escapes from Nash inflation. Am. Econ. Rev. 96 (4), 1193–1224.
Sargent, T.J., 1999. The Conquest of American Inflation. Princeton University Press, Princeton, NJ.
Sargent, T.J., Williams, N., 2005. Impacts of priors on convergence and escapes from Nash inflation. Rev. Econ. Dyn. 8 (2), 360–391.
Shiller, R.J., 1978. Rational expectations and the dynamic structure of macroeconomic models: a critical review. J. Monet. Econ. 4 (1), 1–44.
Slobodyan, S., Wouters, R., 2012. Learning in an estimated medium-scale DSGE model. J. Econ. Dyn. Control 36 (1), 26–46.
Smets, F., Wouters, R., 2007. Shocks and frictions in us business cycles: a Bayesian DSGE approach. Am. Econ. Rev. 97 (3), 586–606.
Stock, J.H., Watson, M.W., 2003, December. Has the business cycle changed and why? NBER Macroeconomics Annual 2002 BST NBER Chapters, vol. 17. National Bureau of Economic Research, Inc., pp. 159–230.
Weber, A., 2010. Heterogeneous Expectations, Learning and European Inflation Dynamics. Cambridge University Press, pp. 261–305 (Chapter 12).
Williams, N., 2003, January. Adaptive Learning and Business Cycles. Mimeo.