UNIVERSITY OF CAMBRIDGE
Faculty of Economics

CAMBRIDGE WORKING PAPERS IN ECONOMICS

Optimal Feasible Expectations in Economics and Finance

Alfred
Lake

Abstract

Trying to estimate rational expectations does not usually minimise forecast error when forecasting macroeconomic or financial variables in reality. This is because, with samples of realistic length, optimal feasible forecasts contain conditional biases that reduce forecast variance. I demonstrate this by using penalised factor models to show that statistically simple inflation forecasts, primarily based on past inflation, are optimal even when other relevant financial and economic variables are available. I also show that US household inflation forecasts display many similarities to these simple optimal forecasts, but also contain mistakes that increase forecast error. Therefore a combination of `optimal feasible expectations' and behavioural errors explain US household inflation forecasts. This suggests that optimal feasible expectations, with additional behavioural errors in some cases, could explain forecast formation across economics and finance.

Reference Details

CWPE	20105
Published	11 November 2020
Key Words	Forecasting, Expectations, Uncertainty, Shrinkage, Ination, Nominal Rigidities, Factor Models
JEL Codes	E37, D84, E70, C53
Website	www.econ.cam.ac.uk/cwpe

Optimal Feasible Expectations in Economics and Finance

Alfred Lake*

November 2020

Abstract

Trying to estimate rational expectations does not usually minimise forecast error when forecasting macroeconomic or financial variables in reality. This is because, with samples of realistic length, optimal feasible forecasts contain conditional biases that reduce forecast variance. I demonstrate this by using penalised factor models to show that statistically simple inflation forecasts, primarily based on past inflation, are optimal even when other relevant financial and economic variables are available. I also show that US household inflation forecasts display many similarities to these simple optimal forecasts, but also contain mistakes that increase forecast error. Therefore a combination of 'optimal feasible expectations' and behavioural errors explain US household inflation forecasts. This suggests that optimal feasible expectations, with additional behavioural errors in some cases, could explain forecast formation across economics and finance.

Keywords: Forecasting, Expectations, Uncertainty, Shrinkage, Inflation, Nominal Rigidities, Factor Models

JEL Classifications: E37, D84, E70, C53

*University of Cambridge: al741@cam.ac.uk. The author is grateful to the Cambridge Trust and Girton College for financial support.

1 Introduction

Understanding inflation expectations and forecasts is central to macroeconomics. Inflation expectations drive inflation itself through wage bargaining and price setting, so affect nominal rigidities and the real impacts of aggregate demand shocks. Therefore the responsiveness of inflation expectations to available information on macroeconomic shocks affects the answer to crucial questions such as the reaction of unemployment to financial crises or the ability of government spending to boost output.

Rational expectations have been the most common approach to modelling how inflation forecasts are formed in academic economics and finance in recent years, although empirical and theoretical work has suggested behavioural alternatives (Coibion et al., 2018). They are defined as agents' expectations being the conditional expectation of future variables, conditioning on available information[1]. Sheffrin (1996) provides a full mathematical definition of this while the original definition is available in Muth (1961). They imply that agents' expectations should react to publicly available information in the same manner as future realised inflation reacts (Lovell, 1986).

Rational expectations will not usually be the optimal feasible expectations for agents when forecasting a variable. If rational expectations are feasible then they will be the optimal feasible expectations for an agent, using mean square forecast error to define optimality (Diebold, 2017). However they will only be feasible if the agent can either deduce or perfectly estimate the relevant parameters of the conditional distribution of the variable being forecasted. Given the complexity of modern economies it is simply not possible to deduce rational expectations without estimation from data in the vast majority of circumstances. Therefore estimation from data must be used to form expectations. This has driven a large learning literature studying whether expectations based on learning from data converge to rational expectations, which is surveyed by Evans and Honkapohja (2012). Since

[1] Full information rational expectations, as described by Coibion et al. (2018) and commonly used in academic macroeconomics, also require that complete knowledge of the economy is available.

papers in this literature are primarily interested in convergence, they often assume agents have access to infinite observations of data. With infinite data an agent can use a conditionally unbiased and consistent estimate as the conditional expectation, such as that formed by approaches similar to regressing realised values of the variable being forecast on all available past information. Such an approach would converge to the conditional expectation, i.e. rational expectations, so rational expectations are feasible with infinite data.

However with a finite series of data rational expectations are not usually feasible, as conditionally unbiased estimators, such as those produced by regressing the variable being forecast on past information, will vary slightly around the conditional expectation as a result of estimation error. Any conditionally biased estimator will also contain clear deviations from rational expectations. Therefore rational expectations will not be the optimal feasible expectations in most settings. It also may not be optimal to try to limit the conditional biases that one imposes in expectations to make estimating them feasible. This is because it may be worth shrinking estimated expectations towards statistically simple expectations to reduce forecast variance, even though this introduces greater conditional biases. This insight is at the heart of modern machine learning (Ahmed et al., 2010) and Bayesian approaches to forecasting (De Mol et al., 2008) and is also present in frequentist forecasting approaches (Bai and Ng, 2008). Therefore the optimal feasible expectations in most forecasting situations in economics and finance will be clearly different from, and statistically simpler, than rational expectations.

I begin this paper by discussing why some shrinkage is likely to be needed when forming expectations for the vast majority of macroeconomic and financial variables, as a result of the large number of potentially relevant data series available relative to the number of observations of each series[2]. I also discuss why it is very often likely that additional shrinkage towards statistically simpler specifications will improve the bias-variance trade-off of forecasts as a result of reducing estimation error and so reduce measures of forecast error. However the precise level of shrinkage in optimal feasible expectations in a particular setting is ultimately an empirical issue, so I then

[2]A phenomenon known as fat data (Koop, 2017)

analyse the empirical importance of this shrinkage when forecasting US inflation.

Specifically I consider adding a number of different potential predictors of inflation to a baseline auto-regressive forecast of inflation. I estimate these forecasts in a number of training sets using weighted ridge specifications with different levels of statistical shrinkage applied to each additional predictor and then take the optimal degree of shrinkage as the one which minimises measures of forecast errors in test sets. The results suggest that a large degree of shrinkage should be applied to most variables[3]; indeed the optimal forecast virtually only uses information on past inflation and components of inflation. These results are closely linked to those from the empirical inflation forecasting literature, which show that univariate inflation forecasts are hard to beat in forecasting horse-races (Stock and Watson, 2008). However, unlike this literature, I also estimate equivalent specifications without shrinkage and find that inflation does have economically and statistically significant associations with some of the predictors. This implies that the high levels of optimal shrinkage do not just come from inflation being uncorrelated with past information, they also come from it being worth conditionally biasing inflation forecasts towards statistically simpler forecasts to reduce conditional forecast error. Therefore the optimal feasible inflation expectations are very different from rational expectations.

Finally I analyse the conditional biases in surveys of actual US household inflation forecasts using an approach similar to that in the existing literature. I find that there are significant biases, particularly in the response to changes in past broad inflation and to financial cycle indicators. Many of these conditional biases appear to arise from using the same conditional biases as estimated optimal feasible expectations, such as the limited response to broad changes in past inflation. However household forecasts are shown not to be the optimal feasible expectations, as they perform worse in pseudo out of sample forecast comparisons than feasible empirical alternatives[4]. Therefore both optimal feasible expectations and behavioural

[3]The results only consider the linear effects of the variables, but given the number of potential non-linear effects far more shrinkage would be needed to even make estimation with a wide range of non-linear effects feasible.

[4]This is unsurprising given the clear evidence that many people have a poor understanding of inflation (Del Giovane et al., 2008)

mistakes are likely to have a role in explaining US household inflation forecasts.

I also suggest optimal feasible expectations as a new general class of expectations, formally defined as the expectations that are predicted to minimise the relevant measure of forecast error out of the set of expectations that it is feasible for agents in the real world to estimate. Optimal feasible expectations are likely to differ materially from rational expectations in most circumstances as they are likely to incorporate conditional biases associated with being statistically simple. Indeed, given the importance of parsimony in forecasting many variables (Kim and Swanson, 2018), despite their no doubt numerous true links to one another, optimal shrinkage is likely to cause optimal feasible expectations to be dramatically different to rational expectations in a large number of macroeconomic settings. I suggest that we should generally conceive of macroeconomic expectations as optimal feasible inflation expectations, with the possible addition of behavioural errors in settings in which agents do not act optimally.

Work on how inflation expectations are formed has a long and important history in the macroeconomic literature that includes the discussions of money illusion in Keynes (1936), the adaptive inflation expectations in Friedman (1977), the model-specific rational price expectations in Lucas (1996) and the behavioural pricing in Akerlof (2002). However the work in this paper is most closely related to, and contributes to, three relatively distinct branches of the existing literature.

Firstly, this paper relates to the literature studying learning and expectations in macroeconomics, as surveyed in Evans and Honkapohja (2012). In this literature work tends to investigate the implications of agents learning expectations from data in theoretical macroeconomic models. As the majority of this literature tends to focus on whether such learning behaviour leads to models converging to rational expectations equilibria, it is common practice to assume that agents have access to an infinite series of relevant data (Evans and Honkapohja, 2012). However, as described above, in this case it is feasible and optimal to use a conditionally unbiased and consistent estimate of rational expectations, such as that given by approaches based on least squares, which then simply implies that agents use rational expectations[5]. When agents have finite data it is not feasible to use rational expectations, as consistent estimators will not converge to the true conditional expectations. However it may be possible to use a conditionally unbiased estimator, such as approaches based on least squares similar to that in Orphanides and Williams (2007), which implies that agents use rational expectations plus noise. However the papers that are most closely related to this paper are those in which agents with finite data use methods that give conditionally biased expectations. For instance Hommes et al. (2019) assume agents use least squares but only applied to an auto-regressive rule while Chung and Xiao (2013) assume that agents use a vector auto-regression with a subset of relevant variables.

The justification for these learning methods is that the authors are looking for a method that balances tractability in the theoretical model considered with being a good approximation for what some forecasters do in practice. I contribute to this literature by studying the optimal expectations that are feasible for an agent to use, rather than the feasible expectations that some agents may use in practice. My empirical approach frees me to do this and has the advantage of allowing me to study shrinkage in the real world. I demonstrate that in the case of US inflation forecasting the optimal feasible expectations contain large conditional biases, conditioning on important macroeconomic series and series that are often used as predictors of inflation. This is because shrinking expectations towards simpler forecasts reduces conditional forecast variance sufficiently to more than offset the conditional forecast bias imparted. I also suggest why similar shrinkage is also likely to be used in the optimal feasible expectations in the vast majority of macroeconomic settings. This is hugely important as it implies that agents learning optimally will use expectations that are often very different from rational expectations, despite this being a fundamental justification of the rational expectations revolution (Coibion et al., 2018). I therefore suggest optimal feasible expectations as a new general class of expectations that are defined as the expectations that have the lowest predictable forecast error

[5]Note I am discussing whether an approach implies that agents use rational expectations from an infinite sample of relevant data, not whether an approach leads to a specific model converging to a rational expectations equilibria of that model.

out of the set of expectations that agents in the real world could actually estimate. These are likely to be conditionally biased towards statistically simple specifications, so will usually be much statistically simpler than rational expectations.

Secondly, this paper relates to the econometric literature on forecasting inflation. There are a very large number of papers that analyse different approaches for forecasting inflation, in terms of method and/or predictive variables, and compare pseudo out of sample forecast error measures. Reviews of this literature are provided for traditional econometric methods in Stock and Watson (2008) and Faust and Wright (2013), while Medeiros et al. (2020) extend this analysis to machine learning methods. A key message that emerges from these reviews is the importance of parsimony. Simple auto-regressive benchmarks forecast extremely well: they are hard to consistently out-perform and effectively impossible to consistently out-perform by a large margin at horizons less than two years. Those methods that do appear to out-perform them minimise and constrain additional estimation. These include factor models with a very limited number of macroeconomic factors (Stock and Watson, 2002), extensions to benchmark models that still only use price data but allow different components of inflation to have different effects (Stock and Watson, 2016) and very heavily pruned random forests that allow some heavily constrained effects of employment variables (Medeiros et al., 2020). Theoretical restrictions derived from DSGE models are not useful for improving forecasting performance (Giacomini, 2015), however central banks targets, or proxies for them, become the optimal forecasts at horizons much beyond two years (Faust and Wright, 2013). This appears sensible, as central banks aim to target inflation in the medium term, however at horizons of two years or less lags in the effects of monetary policy (Havranek and Rusnak, 2013) and central banks' preferences for gradual adjustment of interest rates (Coibion and Gorodnichenko, 2012b) suggest that inflation deviations from targets are forecastable.

This literature currently does not address precisely why it is not optimal to add information on particular variables to auto-regressive benchmarks and I contribute to this literature by studying why this is the case. I initially assess how much shrinkage is optimal to apply to a series of macroeconomic variables, that include the main variables often used in macroeconomic models and variables commonly used in inflation forecasting. In line with the existing literature I find that the majority of variables should have total shrinkage applied to them, implying that one should virtually only use information on price series to form inflation forecasts. This information should not be used naively though, as different types of inflation should be allowed to have effects that differ but are constrained to limit estimation error. However I go on to provide the first comparisons of the shrunken estimates of the association between each variable and future inflation that is optimal for forecasting and consistent OLS estimates of the equivalent actual association. This allows me to analyse whether the high optimal degree of shrinkage comes from the variables simply not having much of an association with future inflation or from the benefits of reducing the variance of the forecast despite this imposing conditional biases because the variables having strong associations with future inflation. The results suggest that for many variables, such as broad inflation and measures of business and financial cycles it is the former, although for variables like wages it is the latter. This is important as it suggests that it is primarily the high degree of uncertainty over the associations between some variables and future inflation that prevents them from being useful in forecasting inflation, rather than the variables simply not having much association with future inflation.

Thirdly, this paper relates to the literature which tests for conditional biases, and hence deviations from rational expectations, in surveys of agents inflation expectations. The main method of testing this in the literature, and the approach used in this paper, is to test whether inflation forecasts and future realised inflation react differently to information that was publicly available at the time of the forecast. Coibion et al. (2018) surveys papers that take this approach. Variables that have been suggested to cause a different response in forecast and realised inflation include lagged forecast errors (Coibion and Gorodnichenko, 2015a), lagged changes in exchange rates (Pesaran and Weale, 2006), narrative shocks (Coibion and Gorodnichenko, 2012a) and lagged energy components of inflation (Coibion and Gorodnichenko, 2015b). Understanding which variables there is a conditionally biased reaction to is important, as this determines which nominal rigidities occur and so helps us to understand how large the nominal rigidities are for the transmission mechanisms of different macroeconomic shocks.

Suggested explanations usually focus on non-optimal behaviour[6], often resulting from some combination of rational inattention or imperfect understandings of the economy (Coibion et al., 2018). This must be at least partly true, as Berge (2018) shows that agent's inflation forecasts can be beaten in pseudo out of sample forecasting by simple auto-regressive moving average models that would have been feasible for agents to use. However it is very important to understand whether some of the specific conditional biases actually arise from optimal feasible behaviour, and so could not be corrected, or if they all arise from potentially correctable behavioural errors. I contribute to this literature by providing what, to my knowledge, is the first evidence on this issue. Using methods similar to the existing literature I estimate the conditional biases in surveys of US household inflation forecasts with respect to a set of macroeconomic variables and show that household forecasts are not optimal feasible expectations as they can be beaten by simple auto-regressive benchmarks. However unlike the existing literature, I then go on to compare the conditional biases in household forecasts to the conditional biases in estimated optimal feasible expectations. I find that the conditional biases in the reaction to many variables, such as broad and narrow inflation, business cycles and exchange rates, are consistent with suggested optimal feasible behaviour. However the reaction to financial cycle information and the amount of noise in household inflation forecasts do not appear to be consistent with optimal feasible behaviour and instead suggest behavioural mistakes. Therefore optimal feasible expectations and behavioural mistakes are each likely to explain part of US households' inflation forecasts.

The rest of the paper is organised as follows. Section 2 lays out my conceptual and econometric framework, Section 3 describes the macroeconomic information used and how I combine some of this information into factors, Section 4 presents the estimates of the conditional biases in optimal feasible inflation expectations, Section

[6]Explanations for some variables, such as aggregate forecast revisions, also include that information on them might not be available in real time, but this is not an issue in this paper as we only consider variables that are publicly available.

5 estimates the conditional biases in surveys of household inflation expectations and compares these to the estimated optimal conditional biases and Section 6 offers some concluding remarks.

2 Conceptual and econometric setup

To clarify the definitions that follow I begin by decomposing future inflation into a component based on public information that is currently available and a component that is unrelated to this information. I then also express forecast inflation in terms of public information that is currently available, as follows:

$$\pi^r_{t+h} = x_t\beta^r + \epsilon_{t+h} \tag{1}$$

$$\pi^f_{t+h} = x_t\beta^f \tag{2}$$

where π^r_{t+h} is inflation at time $t+h$, π^f_{t+h} is an agent's forecast at time t of inflation at time $t+h$, x_t is a vector of information that is publicly available at time t, β^r is a vector of true coefficients, β^f is a vector of coefficients that agents use in their forecasts and ϵ_{t+h} is the component of inflation at time $t+h$ that is unpredictable a time t with public information.

This expression is very general, as x_t could include lagged information or information which is non-linear in underlying indicators. It could also include information that is unrelated to future inflation, so that some of the values in β^r could be zero.

The definitions of the terms used are then as follows. I define the set of feasible expectations as expectations based on choices of β^f that agents can actually use in realistic settings. For instance it would be feasible to use OLS to estimate the values based on past observations. It would also be feasible to choose to set the value on lagged inflation to one and all other values to zero. I define optimal feasible expectations as the specific expectations in the set of feasible expectations that ex ante can be predicted to minimise the relevant measure of out of sample forecast error. Rational expectations are defined following Sheffrin (1996), and originally Muth (1961), as expectations that are equal to the true conditional expectation of future variables, conditioning on available information. Applying this definition in this settings yields that rational expectations are the expectations given when $\beta^f = \beta^r$.

I now consider whether rational expectations will be the optimal feasible expectations in realistic settings. First it is important to note that if rational expectations are feasible then they will be optimal, as defined by the mean square forecast error[7], since the conditional expectation statistically minimises mean squared forecast error (Granger and Newbold, 1986). If an agent had infinite relevant data to learn from then they could use any consistent estimator of β^r to obtain an estimate essentially equal to β^r that could then be used to construct rational expectations[8]. For instance one could use past observations to estimate Equation 1 using OLS with all potential predictors of inflation in x_t to obtain an estimate of β^r that is statistically perfect. The agent could then use this perfect estimate of β^r as β^f, so rational expectations are feasible in this scenario and hence they are also the optimal feasible expectation.

However in reality, agents clearly only have a finite sample of data available to them. Forecasts often need to be constructed at horizons of at least a year, however samples of relevant data are usually short relative to these horizons and will not necessarily increase over time, as economies experience huge structural changes that decrease the relevance of older data. For instance, formal tests (Stock and Watson, 1996) and institutional change suggest that the economic dynamics of countries now are very different from the dynamics in the period before the 1980s, when most policymakers were fully Keynesian and the internet had not yet been invented. They are likely to be even more different to the dynamics from earlier periods when many of these countries engaged in active global wars with one another. Therefore data from previous structural eras is unlikely to be of significant quantitative relevance for an agent seriously engaged in inflation forecasting (Stock and Watson, 2008). There are also strong reasons to believe that this phenomenon will continue in the

[7]A single point forecast can only generally minimise a single forecast accuracy measure and the mean square forecast error is one of the most common measures Diebold (2017).

[8]Technically this applies to stationary variables. One would need to difference non-stationary variables until stationarity was achieved before applying this process. Then the results of this process and the current values of the variables could then be used to construct rational expectations.

future. For instance it seems extremely likely that there will be significant structural economic change as a result of climate change, artificial intelligence and the growth of countries like China.

In reality there are huge number of potential predictors that are likely to have some effects on inflation relative to samples of data of these lengths[9], as any variable that affects how firms set prices will have some effect on future inflation at shorter horizons. Combining similar variables may reduce the number of series that could be used but lags and non-linear transformations will increase this number and it will remain very large in practice. For instance Refinitiv Datastream and similar services provides millions of macroeconomic data series yet even samples dating to World War 2 only contains hundreds of months of observations. Therefore using conditionally unbiased approaches is simply not feasible. For instance, OLS estimates of Equation 1 cannot be estimated while including many of the macroeconomic series that are available. Therefore agents will generally need to use an estimation approach that shrinks forecasts, partially or even absolutely, towards statistically simpler specifications for estimation to be feasible. This implies that in practice all feasible expectations are likely to contain conditional biases, so rational expectations will not be feasible.

Even if one had incorporated enough shrinkage to make estimation feasible it may well be optimal to include more shrinkage. The optimal feasible approach needs to optimally balance conditional forecast bias against conditional forecast variance, conditioning on the information available. This can be seen most clearly when using the mean squared forecast error as the measure of forecast performance. Consider the following decomposition of the mean squared forecast error, where all expectations are conditional on the information in x_t and the decomposition uses Equation 1, into the components that contribute to it:

[9] A phenomenon that has been more broadly been described as big data in macroeconomics being 'fat' data, with many series but relatively few observations of each series (Koop, 2017)

$$
\begin{aligned}
MSFE &= E\,(\pi^f_{t+h} - \pi^r_{t+h})^2 \\
&= E\,(\epsilon_{t+h})^2 + E\,(x_t\beta^f - x_t\beta^r)^2 - 2E\,(\epsilon_{t+h}(x_t\beta^f - x_t\beta^r)) \\
&= E\,(\epsilon_{t+h})^2 + E\,(x_t\beta^f - E\,(x_t\beta^f) + E\,(x_t\beta^f) - x_t\beta^r)^2 \\
&= E\,(\epsilon_{t+h})^2 + E\,(x_t\beta^f - E\,(x_t\beta^f))^2 + (E\,(x_t\beta^f) - x_t\beta^r)^2 \\
\\
&= unpredictable\ component\ +\ forecast's\ variance\ +\ (forecast's\ bias)^2
\end{aligned} \tag{3}
$$

The choice of the parameters, β^f cannot change the unpredictable component but they will affect the conditional variance and the conditional bias. An approach that is just feasible, such as using OLS estimates of Equation 1 with as many series in x_t as observations may minimise conditional biases, but is also very likely to impart a large amount of estimation error that contributes to conditional variance. Whereas using an approach that did not involve estimation, such as assuming a random walk, would minimise conditional variance but is very likely to impart conditional bias. Therefore there is typically a bias-variance trade-off to consider in the choice of how much shrinkage an agent should use when choosing β^f.

The statistically simple specifications that it is optimal to shrink forecasts towards will not usually be given by theoretical macroeconomic models. On a purely empirical level this is currently true, as the literature survey in Giacomini (2015) shows that the full results of quantitative macroeconomic models are not useful for improving the forecasts of typical macroeconomic variables given by purely statistical approaches. Giacomini (2015) suggests that the limited results to the contrary are a product of the data mining that is fundamental in creating a theoretical model of an economy based on recent experience and then testing its ability to forecast in a sample that includes the periods on which recent experience is based. On a more fundamental level it is likely to continue to be true as theoretical macroeconomic models usually only offer predictions conditional on structural shocks and state variables that are not well observed in practice (Chung and Xiao, 2013), so proxies for them may not have the predicted effects.

There are a limited number of cases where useful guesses of coefficients in β^f can be deduced without data[10], some of which are discussed in Giacomini (2015). In a very limited number of cases these may even allow expectations that are close to rational to be used: for instance heavily shrinking long-term inflation forecasts in some countries towards the countries inflation target. However usually this information will be very imprecise and so will not allow rational expectations to be used, for instance one may know that the Covid-19 pandemic and climate change will reduce future global output but there is enormous uncertainty about the magnitude of the reductions. Therefore in the the vast majority of economic and financial situations the natural choice is to shrink the vast majority of coefficients in β^f towards zero. The optimal degree of shrinkage can then be based on a combination of how relevant an agent thinks a variable is likely to be, for instance ruling out variables that are likely to have little association with the macroeconomy in question so are unlikely to have large effects, and empirical methods, such as pseudo out of sample tests or Bayesian model averaging.

I therefore suggest a new class of expectations: optimal feasible expectations. These are formally defined as the point expectations that are ex ante predicted to minimise the relevant measure of forecast error out of the set of expectations that it is feasible for agents to use in practice[11]. Based on the above discussion I suggest that in the vast majority of macroeconomic settings optimal feasible expectations are likely to be statistically simpler than rational expectations, as many variables effects will be shrunk significantly towards zero, so they will generally incorporate conditional biases. However the exact size of the conditional biases in optimal feasible expectations, and hence their differences with rational expectations, is an empirical question. It depends on the degree of shrinkage that is optimal to apply

[10]However shrinking coefficient towards these values may actually increase the degree of shrinkage in optimal feasible expectations relative to shrinking them towards zero, as the same reduction in conditional forecast variance from absolute shrinkage could then be achieved with less conditional bias.

[11]Optimal feasible expectations could therefore vary for different agents if they aim to minimise sufficiently different measures of forecast error in the same setting. However this is partly a product of analysing point forecasts and is not the focus of this paper, so is not explored here.

to variables that have large associations with future inflation. I therefore now turn to examining the degree of optimal shrinkage to apply to variables in US inflation forecasting. The specific variables I choose are ones that are thought to transmit shocks to inflation in many macroeconomic models and so have long been suggested in the literature as potentially having associations with future inflation.

It is worth noting however that even if empirically observed shrinkage is relatively small this could imply large deviations from rational expectations equilibria, as it may represent the endpoint of a feedback loop. For instance, consider agents applying shrinkage with regards to information on a macroeconomic shock, so that their expectations responded less than rational expectations would to information on the shock. This could in turn reduce the response of realised inflation itself to the shock, relative to rational expectations equilibria, which could lead to even greater differences between expectations and those in rational expectations equilibria, creating a feedback loop. Therefore actual data may be generated by the end point of such a feedback loop and relatively limited empirical shrinkage could still imply large nominal rigidities relative to comparable rational expectations equilibria.

Since estimating without shrinkage is infeasible in reality, as discussed above, I start with an extremely parsimonious specification and then consider how much it is worth shrinking the effects of additional macroeconomic variables that are added to this benchmark[12]. As well as estimating the degree of shrinkage that is optimal to apply to the predictive associations of each of these variables when forecasting future inflation, I also estimate the true association between each variable and future inflation. This lets me analyse the size of the conditional biases present in the estimated optimal feasible inflation expectations.

The baseline specification that I start with is an extremely simple direct autoregressive model estimated by OLS. It simply expresses inflation at time $t+h$, π^r_{t+h}, in terms of inflation at time t, π^r_t, and a constant:

$$\pi^r_{t+h} = \gamma_0 + \gamma_1 \pi^r_t + \varepsilon_{t+h} \tag{4}$$

[12]This also seems sensible given the importance placed on extreme parsimony by the inflation forecasting literature discussed in Section 1.

I can then consider the optimal level of shrinkage to apply to additional macroeconomic variables[13] using pseudo out of sample inflation forecasting performance. To do this I take many overlapping sub-samples from my sample and then in each of these training sub-samples calculate estimates of the coefficients on these additional variables with different levels of shrinkage. The optimal level of shrinkage can then be taken as the one which minimises the pseudo out of sample forecasting error from the remaining test datasets. The approach is therefore a conservative one for estimating the optimal degree of shrinkage, as a new test set is not used for every variable. This pseudo out of sample approach is a common method of setting the level of shrinkage in machine learning approaches such as those in Medeiros et al. (2020). Note that the optimal level of shrinkage will become very small as the sample becomes very large, so this approach can still produce a consistent forecast.

I implement the shrunken estimates using weighted ridge regression[14], which can shrink different coefficients by different quantities and can be expressed as a linear transformation of OLS so can be calculated analytically. I only apply shrinkage to the additional variable that is included. Weighted ridge regression minimises a loss function which combines the OLS loss function with a quadratic penalisation term, so the WR loss function and the OLS loss function can be expressed as follows:

$$Loss\ Function^{WR} = (\Pi - X\beta)'(\Pi - X\beta) + \beta'\Lambda\beta \tag{5}$$

$$Loss\ Function^{OLS} = (\Pi - X\beta)'(\Pi - X\beta) \tag{6}$$

where Λ is a diagonal shrinkage matrix in which the values corresponding to the constant and lagged inflation are zero while the value corresponding to the additional

[13]These variables are only included in linear form, which is conservative as there are so many potential non-linear transformations of variables that attempting to include all of them would require the use of significant shrinkage for estimation to be feasible.

[14]Given the maximum number of variables considered is low, there is little difference between the forecasts produced with this method and alternatives such as lasso or elastic net shrinkage. However there is an analytical solution for weighted ridge, unlike for lasso or elastic net penalisation, making the bootstrapping process used dramatically faster.

variable considered is positive or zero, Π is the vector formed by stacking the dependent realised inflation variable, X is the matrix formed by stacking independent variables and β is the vector formed by stacking coefficients.

Therefore, for each additional variable considered, I estimate the following specification by OLS for the whole period and by a series of weighted ridges with multiple levels of shrinkage over each of a series of training periods for each additional variable v^i:

$$\pi^r_{t+h} = \gamma_0 + \gamma_1 \pi^r_t + \alpha v^i_t + \varepsilon_{t+h} \tag{7}$$

In all specifications I shrink the coefficient on the additional variable included towards zero as a neutral choice and apply no shrinkage to the mean and autoregressive term. The out of sample forecasting results with different levels of shrunken coefficients then provide estimates of the optimal level of shrinkage to be applied to different key variables. Comparing these optimal shrunken coefficients to the OLS coefficients is then an estimate of the conditional biases imposed on the information contained in these variables[15]. The greater the difference between the optimal shrunken coefficients and the OLS coefficients the greater the conditional biases in estimated optimal feasible expectations. Larger conditional biases imply larger deviations of optimal feasible expectations from rational expectations and so larger nominal rigidities that arise from inflation expectations.

The deviation of optimal feasible expectations from rational expectations implies that any conditional biases in agent's forecasts are not necessarily a deviation from optimal behaviour in the real world. Therefore in the second part of my analysis I estimate if there are conditional biases in household inflation forecasts with respect to the variables considered above and compare these conditional biases to those in the estimated optimal feasible expectations. To do this, I compare the OLS and weighted ridge estimates from the previous specification with OLS estimates of the equivalent specification with household inflation forecasts as the dependent variable as follows:

[15]Note they will include both the direct information included in the variable itself and the information included through its correlations with all other variables that have not been included.

$$\pi^{f}_{t+h} = \gamma_0 + \gamma_1 \pi^{r}_{t} + \alpha v^{i}_{t} + \varepsilon_{t+h} \tag{8}$$

Differences between the estimated OLS coefficients with realised inflation and household inflation forecasts as the dependent variables imply that there are conditional biases in household forecasts[16], so household forecasts deviate from rational expectations. Note that this is true even if one includes a subset of the data available to agents (Sheffrin, 1996). I formally test the differences between these coefficients from Equations 7 and 8 and obtain confidence intervals for the difference using a joint block-bootstrap. Specifically, I use a bias-corrected version of Hall's empirical bootstrap approach, which allows for auto-correlated errors and parameter distributions which are skewed and incorrectly centered.

I also compare the similarities between the OLS coefficients with household inflation forecasts as the dependent variables and the estimated optimal weighted ridge coefficients. This is because similarities suggest that the conditional biases considered are consistent with the conditional biases in estimated optimal feasible expectations, whereas differences suggest they are not. Finally, I also check whether households' forecasts are consistent with being optimal feasible expectations by comparing their out of sample forecast performance with that given by my parsimonious benchmark, as this benchmark is feasible and approaches like this have long been known to forecast reasonably well (Gordon, 1982). If the household forecast performance is as good as or better than the estimated forecast performance of this benchmark then this is consistent with households using optimal feasible expectations. Although one should remember that there clearly may be better feasible alternatives to my benchmark available, so this a necessary and not sufficient condition. However if households' forecast perform worse than my benchmark then this strongly implies that households make behavioural mistakes that cause their

[16]It is theoretically possible that 'peso problems' could explain such differences in short samples, however this seems unlikely to be important in a sample which includes the financial crisis, the dot-com bubble and many other extreme events. Additionally the effects of large infrequent events seem especially unlikely to be something that households could estimate perfectly and so incorporate in line with rational expectations.

expectations to deviate from optimal feasible expectations.

3 Macroeconomic data and factors

My two primary dependent variables are household inflation forecasts and realised inflation. The household inflation forecasts are taken from the Michigan Survey of Consumers: they are one year ahead inflation forecasts and the questionnaire aims for quantitative responses with prompts provided if necessary[17]. I choose an annual horizon as this is long enough for many shocks to have some inflationary effects, but is not long enough for the Federal Reserve to have resolved these effects, due to lags in the effect of monetary policy (Havranek and Rusnak, 2013) and a preference for gradual monetary policy action (Coibion and Gorodnichenko, 2012b). The household forecasts are usually based on a sample of approximately 500 people, of which up to 20% give non-quantitative answers. These non-quantitative answers are hard to reconcile with rational expectations or optimal feasible expectations, so strongly suggest that household forecasts may not be optimal even before any formal analysis is conducted.

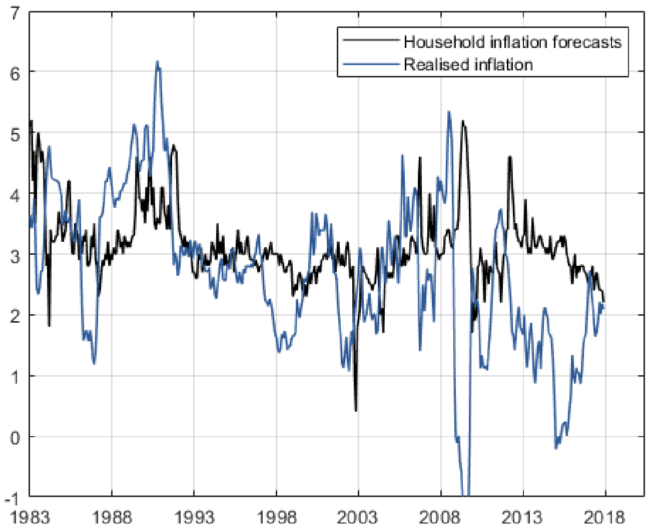
I take the consumer price index as my measure of realised inflation. This is because its definition is methodologically most suitable, as it aims to capture the experienced inflation of consumers, which is not true of alternatives such as the personalised consumption expenditures index. Its mean level is also closer to the mean household inflation forecast than the mean level of alternatives such as the personalised consumption expenditures index, which supports it being the appropriate inflation index. Inflation is widely considered to be stationary in the absence of structural breaks, as it does not seem feasible that the central bank would allow significant deviations from its goals and its long-term mean has remained similar over recent decades. Breitung and Eickmeier (2011) find that there is a structural break in inflation at the start of Paul Volcker's chairmanship of the Federal Reserve and several of the variables in my sample are only available from close to this point onwards, so I begin my sample near this point. If there is any additional structural

[17] The specific questionnaire can be accessed online through http://www.sca.isr.umich.edu/

change over time this should be picked up by the inflation factor, so will not cause spurious results.

The main sample for inflation as a dependent variable therefore runs from the annual price growth up to January 1983 to the annual price growth up to December 2017. The sample for dependent inflation expectations necessarily covers the expectations for the same period and the sample for control variables is lagged by a year.

Figure 1: Realised inflation and household inflation forecasts



Notes: Plots of the median household annual inflation forecast for the past year and annual consumer price index inflation. The vertical axis is in percentage points and the horizontal axis is in years.

Figure 1 plots household inflation forecasts and realised inflation over the sample. Household forecasts are generally of a similar approximate level to realised inflation, however there are several features which may seem surprising if one were expecting inflation forecasts to be formed by rational expectations. Spikes in household forecasts often follow, instead of precede, spikes in realised inflation and there are long periods of divergence between forecasts and realised inflation. These features suggest that inflation expectations are not formed rationally, although this will be examined in much more detail in Section 5.

The additional macroeconomic variables I consider adding to the forecasting procedure include some of the most important potential transmitters of shocks to inflation and a narrative measure of aggregate demand shocks available in real time. These potential predictors include six series: corresponding to business cycles, financial cycles, broad inflation, wages, exchange rates and real-time narrative monetary shocks. They therefore include equivalents of the macroeconomic variables suggested as potential predictors of inflation in Stock and Watson (2008) as well as one of very few narrative measures of shocks available in real time. I do not include the measures from specific financial markets that Stock and Watson (2008) suggest including as proxies for expectations themselves, as in this paper the expectations being formed are viewed as the dependent variable to be explained, so including proxies for them as an independent variable would not be helpful. In all cases I take the variables from the Fred MD database or the broader FRED database except for the monetary shocks which are taken from Gertler and Karadi (2015).

In the case of the first three series (business cycles, financial cycles and broad inflation) many monthly measures are available so I combine them using a factor approach, whereas for the latter three series (wages, exchange rates and real-time narrative monetary shocks) there are few series available so I simply use the corresponding series in Fred MD or Gertler and Karadi (2015). I produce the factors using the principal components approach of Stock and Watson (2002). This is applied separately to different groups of variables, so each variable only loads on one factor, as suggested by Bernanke et al. (2005). This ensures statistical identification and also gives each factor a clear economic interpretation. Bai and Ng (2006) show that the factors converge at rate $min(N,T)$, whereas if the factors were known then the coefficients would converge at rate $\sqrt{T}$. Therefore it is a reasonable approximation to treat the factors as known if N is reasonably large compared to $\sqrt{T}$, which is the case here. Indeed, it may well still be an improvement over using specific variables to proxy for each factor, which would remove the estimation issue but potentially introduce significant measurement error.

I take the majority of the variables from the Fred MD database. For the business cycle factor I take 15 variables from the output and income section and 21 variables from the labour market section, which are all in real terms. For the price factor I take 19 variables from the prices section and add 16 extra price series from the broader FRED database. For the financial cycle factor I take 7 credit series from the money and credit section and add 2 extra credit series and 31 house price series from the broader FRED database. For the exchange rate series I take the trade weighted US Dollar index against major currencies, where a rise implies an appreciation of the dollar, and for the wage series I take the average hourly earnings of goods producing workers. This gives 36 business cycle series, 35 inflation series and 40 financial cycle variables in a sample where $\sqrt{T}$ is approximately 20. Therefore in each case $\sqrt{T}/N$ is small, so any estimation error in the factors will be limited relative to estimation error of the coefficients.

The narrative monetary shocks are taken from Gertler and Karadi (2015). They are constructed as the high frequency changes in federal funds futures markets around federal reserve announcements and are discussed and contrasted to other shocks in Ramey (2016). It is important to note that I do not necessarily give the narrative shocks a causal interpretation, as Miranda-Agrippino (2016) shows that they respond to Federal Reserve forecasts. I instead simply view them as one of the most widely-used and reliable measures of monetary shocks available in real time. The sample period is shorter when monetary policy shocks are used, as data is not available for the earlier part of the sample and not usable for the latter part of the sample due to the zero lower bound (Gertler and Karadi, 2015). The sample for inflation as a dependent variable when monetary shocks are used runs from the annual price growth up to July 1990 to the annual price growth up to June 2012. The sample for dependent inflation expectations necessarily covers the expectations for the same period and the sample for control variables is lagged by a year.

All variables are transformed to stationarity, which is primarily by using the FRED MD recommended transformations expressed in annual terms. However inflation is considered stationary over my sample, as discussed above, since it is shorter than the FRED MD sample, so I do not take the second difference of nominal series. All the variables used in the factors are normalised to have zero mean and unit variance before factors are extracted from them. While this uses data from the whole sample in a forecasting exercise it does not change any variable substantially, but just makes them easier to combine and compare. They are also normalised to load positively on a measure of employment, inflation and house prices respectively. I also transform all six additional series that shrinkage is applied to so they have zero mean and unit variance for comparability. A full table of the variables used and the transformations applied is available in the Appendix.

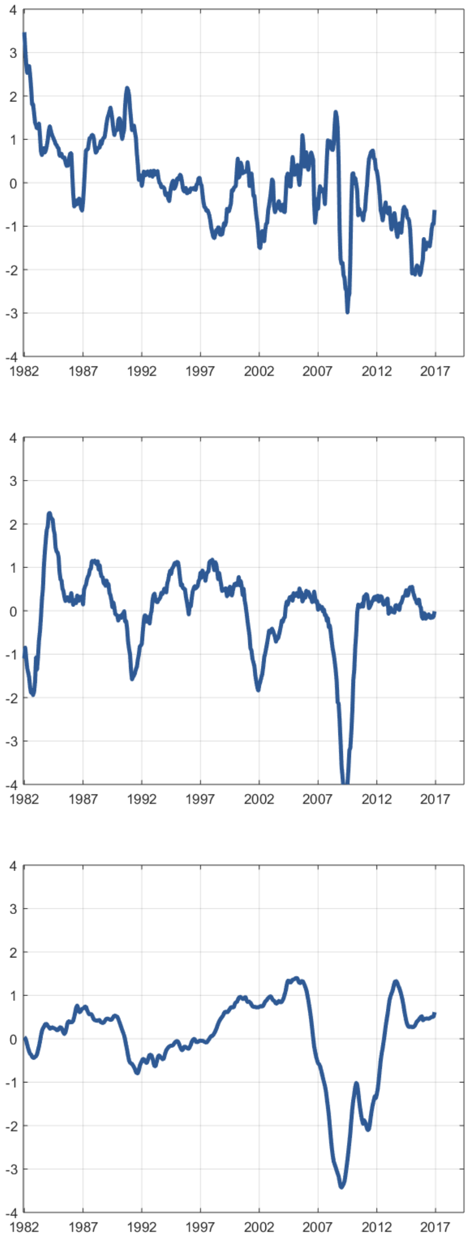
Figure 2 plots the inflation, business cycle and financial cycle factor over the sample. For the inflation cycle, the early parts of the sample contains the large effects of the supply side crises which the Federal Reserve was starting to control, such as the oil price surge at the end of the 1970s. The latter part of the sample has more short-term volatility, although one can see the dip associated with the aftermath of the financial crisis and the subsequent dip associated with the global economic slowdown in 2015 to 2016. The four recessions in the sample are all clearly visible in the business cycle factor and are marked by increases in growth in the recovery after each one. Indeed, this factor could proxy fairly well for the NBER business cycle dating. The financial cycle factor is loosely similar, however the effects of the first two recessions are small and the third is virtually absent, whereas the latter part of the sample is dominated by the huge effects associated with the global financial crisis.

The three factors all load sensibly on their underlying features. In fact, every single variable loads on its factor with the expected sign: all positive for the inflation factor, all positive for the financial cycle factor, negative for the unemployment series and positive for all other series for the business cycle factor. The magnitudes of the factor loadings are also sensible: most are between 0.2 and 0.8 and none are dramatically outside this range[18]. Therefore the factors appear to capture the information in the inflation, business cycle and financial cycle indicators well.

Exchange rates, wages and narrative monetary shocks are plotted in Figure 3. There are few clear patterns in the exchange rate graph, as its movements are quite

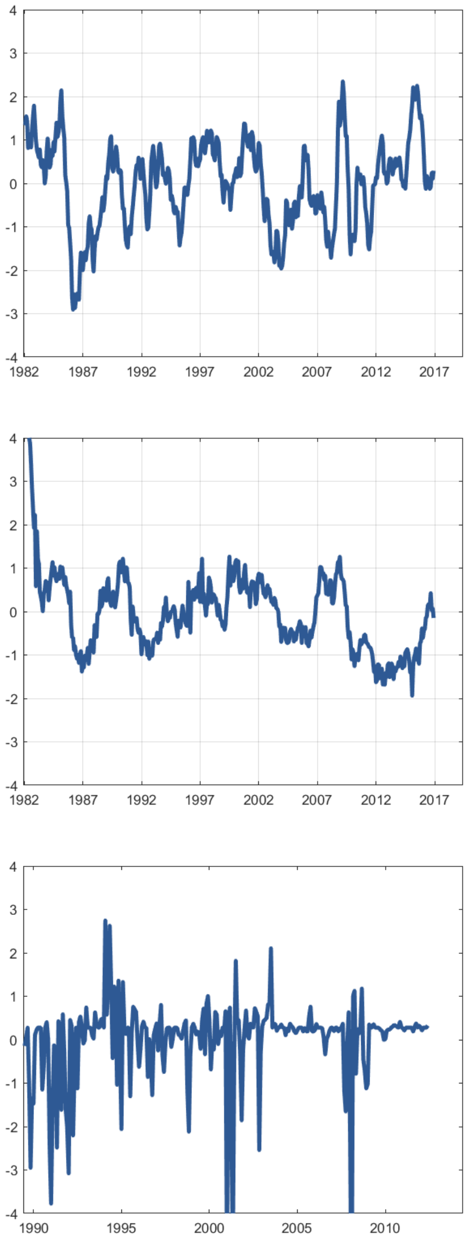
[18]Note that this does not imply that the underlying variables move less in absolute terms than the factor, as they have been normalised to have unit variance.

Figure 2: Inflation factor, business cycle factor and financial cycle factor



Notes: Plots of the factors extracted with the principal components method from transformed data. The vertical axis is in units and the horizontal axis is in years. Since the underlying series are transformed to have zero mean and unit variance and most factor loadings are between 0.2 and 0.8, a one unit change in each factor causes changes in most of its underlying series of between 0.2 and 0.8 standard deviations.

Figure 3: Exchange rates, wages and narrative futures markets monetary shocks



Notes: Plots of the transformed trade weighted exchange rate index, transformed average hourly earnings and transformed narrative futures markets monetary shocks. The vertical axis is in standard deviations of each variable units and the horizontal axis is in years.

volatile. However one can pick out certain large movements, such as the large depreciation in the latter part of the 1980s following the Plaza Accord and the large appreciation in 2014/2015 following monetary divergence between the Federal Reserve and many other developed market central banks. The wage series is also relatively volatile, although one can again notice several large movements, such as the very high wages at the start of the sample as the inflation-wage spiral was being brought under control and the large and sustained declines in wages that occurred in the period following the global financial crisis. The narrative futures market monetary shocks series is the most volatile of all. However one can see especially high volatility in the earlier part of the sample, as well as in the period around the 9/11 attacks and in the period around the global financial crisis.

4 Conditional biases in optimal feasible inflation expectations

I now turn to estimating the optimal degree of shrinkage to apply to each additional variable in inflation forecasts. As discussed in Section 2, I do this using pseudo out of sample inflation forecasting performance. For each variable considered I take many overlapping sub-samples from my sample and then in each of these training sub-samples calculate estimates of Equation 7 with many different levels of shrinkage. I only apply shrinkage to the additional variable added to the baseline specification in each case, so the most shrunken specification corresponds to OLS estimation of the auto-regressive specification in Equation 4 while the least shrunken case corresponds to OLS estimation of the specification in Equation 7, with other levels of shrinkage giving estimates between the two. The estimated optimal level of shrinkage to apply to each variable can then be taken as the one which minimises measures of pseudo out of sample forecasting error from the remaining test data. The training sample sizes are set to 70% of the total sample size in the baseline case, which is relatively typical[19] and ensures the financial crisis period can be in both types of sub-sample. However robustness checks based on increasing or decreasing this sample size are available in the Appendix and the results do not change dramatically in either case.

[19] This gives approximately the same probability of any one observation being in the sample as would be the case if one took a sample with replacement of the same size as the original sample.

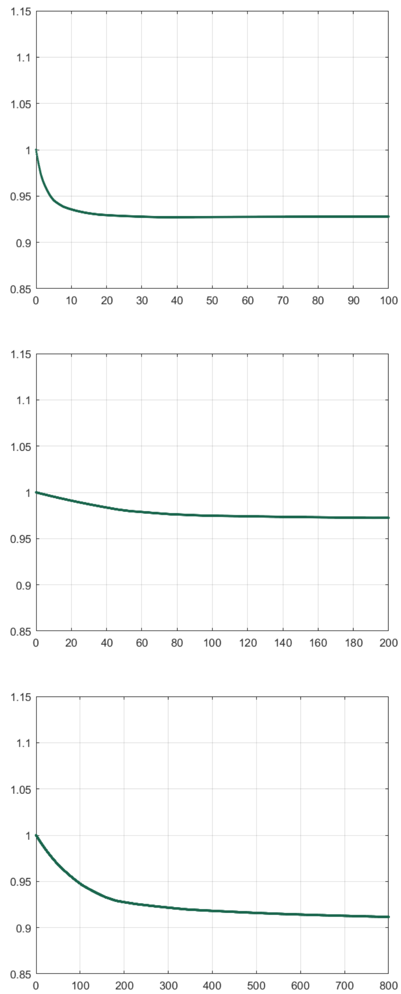
Figures 4 and 5 plot the out of sample forecast performance for different levels of shrinkage applied to each variable. The forecast performance measure used is mean absolute forecast error, although analysis which uses the mean square forecast error is also available in the Appendix and is very similar[20]. The forecast performance is expressed relative to the forecast error using no shrinkage, i.e. that obtained using OLS with the additional variable in question. Therefore values lower than one imply superior performance and values above one imply inferior performance. The optimum shrunken value of each coefficient is reported and compared to the OLS estimate of each coefficient in Table 1.

The first variable I consider is the inflation factor, which captures simultaneous changes in a broad range of the price components of inflation. Therefore including information on this variable allows broad price changes, such as caused by rising consumer confidence, to have different effects on the forecasts produced than the effects of a change in inflation driven by large changes in a small number of prices, such as change in the price of food or oil. The results in the top graph of Figure 4 make it clear that using some shrinkage improves forecast performance: it can reduce the forecast error measure by over 5%. The figures in Table 1 actually show that partial shrinkage is optimal: this is also plotted in the top graph of Figure 4 but is hard to see clearly. This implies that the optimal coefficient on broad inflation when forecasting should be lower than the OLS coefficient of its association with future inflation, but should not necessarily to zero. Although using complete shrinkage and setting the coefficient equal to zero barely reduces the forecast performance from its optimal level. The OLS estimates of the true association between broad inflation and future inflation shows is positive and both economically and statistically significant. Therefore this strongly suggests that optimal feasible expectations should incorporate large conditional biases with respect to information of broad vs narrow inflation, as a result of shrinkage.

The next variables I consider are the business cycle and financial cycle factors. These first of these captures the movements of a set of macroeconomic indicators

[20]These are chosen as they are two of the most common forecast measures (Diebold, 2017) and because the mean square forecast error is the equivalent measure for rational expectations.

Figure 4: Relative forecast error from shrinking information on inflation cycles, business cycles and financial cycles



Notes: Plots of the out of sample mean absolute forecast error of the shrunken estimates of Equation 7 with an additional variable included, presented relative to the out of sample mean absolute forecast error of the equivalent OLS estimate of Equation 7. The inflation factor (top), the business cycle factor (middle) and the financial cycle factor (bottom) are considered. Estimates are based on training sample of 70% of the total dataset. The vertical axis is in relative units, so higher values imply worse performance relative to to the OLS case. The horizontal axis is in values of λ, where higher values of λ imply more shrinkage. When $\lambda \to \infty$ the specification tends to Equation 4.

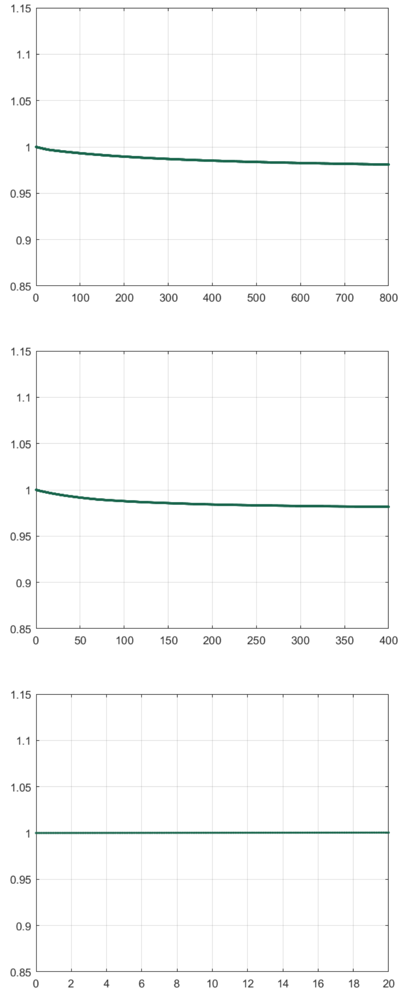
while the second captures the movements of longer term variables in financial markets. The results in the bottom two graphs in Figure 4 and in Table 1 suggest that absolute shrinkage should be applied to these factors, as this reduces the forecast error measure by around 3% and 8% respectively, so optimal feasible expectations should not incorporate information on these variables at all. This suggests that optimal feasible expectations incorporate conditional biases with respect to information on business and financial cycles, as the OLS estimates of their associations with future inflation are positive and economically meaningful, albeit just short of statistical significance.

The variables considered in the upper two graphs in Figure 5 and also in Table 1 are the exchange rate and wage series, which are changes in the trade weighted value of the dollar and hourly earnings respectively. The results again indicate that absolute shrinkage should be applied to these series, as this reduces the forecast error measure meaningfully, so optimal feasible expectations should not incorporate them at all. In the case of wages this appears to be because they have little association with future inflation[21]. However exchange rates have a negative and economically meaningful association with future inflation that is just short of statistical significance, suggesting that optimal feasible expectations contain conditional biases with respect to exchange rate information.

Finally I consider the narrative monetary shock measure taken from federal funds futures markets, which is only available over a shorter sample. The results for this series in the bottom graph of Figure 5 and in Table 1 indicate that it is technically optimal not to apply shrinkage to the effects of this series, so optimal feasible expectations could include this information. However the results also show that the associations of these shocks with future inflation is very small, so that using information on the shocks does not significantly change the forecasts produced and has hardly any effect on forecast performance. Indeed any level of shrinkage produces forecasts with values of the forecast error measure within 1% of each other. It is also interesting to note that the association between the shocks and future inflation is

[21] A result which holds even if one removes the high initial values of the wage series at the very start of the sample.

Figure 5: Relative forecast error from shrinking information on exchange rates, wages and narrative federal funds market monetary shocks



Notes: Plots of the out of sample mean absolute forecast error of the shrunken estimates of Equation 7 with an additional variable included, presented relative to the out of sample mean absolute forecast error of the equivalent OLS estimate of Equation 7. Exchange rates (top), wage (middle) and monetary shocks (bottom) are considered. Estimates are based on training sample of 70% of the total dataset. The vertical axis is in relative units, so higher values imply worse performance relative to to the OLS case. The horizontal axis is in values of λ, where higher values of λ imply more shrinkage. When $\lambda \to \infty$ the specification tends to Equation 4.

Table 1: Estimated true association and optimal associations for forecasting between variables and future inflation

	OLS	Optimal WR
γ_0	1.91*	
	(1.32 to 2.46)	
γ_1	0.27*	
	(0.06 to 0.50)	
α^{inf}	2.08*	0.69
	(1.28 to 2.80)	*(0.05 to 1.42)*
α^{bc}	0.24	0.00
	(-0.07 to 0.50)	*(0.00 to 0.21)*
α^{fc}	0.30	0.00
	(-0.07 to 0.69)	*(0.00 to 0.17)*
α^{er}	-0.18	0.00
	(-0.44 to 0.03)	*(-0.13 to 0.00)*
α^{w}	0.07	0.00
	(-0.26 to 0.35)	*(0.00 to 0.06)*
α^{nms}	0.03	0.03
	(-0.11 to 0.21)	*(0.00 to 0.03)*

Notes: Column 1 shows the OLS estimates of the coefficients from Equation 4 for the baseline variables and the OLS estimates of the coefficient on each new variable from its version of Equation 7. 95% confidence intervals are displayed in standard text beneath OLS estimates. * = statistically significant at the 10% level. Column 2 shows the WR estimate of the coefficient on each new variable from its version of Equation 7 that minimises out of sample error. Bands of shrunken estimates where the forecast error is within 1% of the optimal forecast error are displayed in italicized text beneath WR estimates. By definition these bands are always non-negative or non-positive.

positive, not negative as typically suggested by theory. This may be a result of the fact that these series may capture signals of the Federal Reserve's private economic information as much as they capture true monetary shocks (Miranda-Agrippino, 2016). Therefore this does not appear to suggest that optimal feasible expectations respond to true monetary shocks.

These results have important consequences. Firstly consider how optimal feasible expectations respond to macroeconomic shocks: in particular consider the case of a contractionary monetary shock, i.e. an exogenous increase in interest rates. The response to narrative monetary shocks would imply that inflation expectations would initially rise and there would be no reaction to any change in exchange rates. Inflation expectations would then also not respond to any change in financial cycle and business cycle variables as a result of the shock. They would only start to fall after inflation itself had fallen and even then this response would still be constrained. If higher inflation expectations cause higher future inflation, as seems very likely, then this strongly suggests that optimal feasible expectations would cause large nominal rigidites in the response to such shocks.

Secondly these results provide evidence that optimal feasible inflation expectations contain large conditional biases with respect to some of the most important variables in many macroeconomic models. Therefore they suggest that agents who learn optimally from data will use expectations of key macroeconomic variables that are very different from rational expectations. This undermines one of the major arguments used to try and justify the rational expectations revolution and suggests that macroeconomic models based on rational expectations may be seriously misspecified.

5 Conditional biases in household inflation forecasts

I now turn to estimating the condition biases in household inflation forecasts and assessing whether these are similar to those in optimal feasible inflation expectations. As discussed in Section 2, testing for conditional biases with respect to a given variable is achieved by testing if there are significant differences in the OLS estimates of coefficients from Equation 7 and 8 with that variable is included. Any significant differences would imply different systematic reactions of forecast and realised inflation to the variable and so would suggest conditional biases and hence deviations of household forecasts from rational expectations. I then also analyse whether these conditional biases are the same as those in estimated optimal feasible inflation expectations from the previous section and hence whether household forecasts appear to be similar to estimated optimal feasible expectations.

As discussed in Section 2, I estimate Equations 7 and 8 multiple times, once with each of the six additional variables as v_t^i and once in the baseline case without v_t^i. In all cases I also calculate the difference between each equivalent coefficient from Equation 7 and Equation 8 and bootstrap confidence intervals. Table 2 shows the abridged results of this analysis. The left column shows the results with realised future inflation as the dependent variable, the middle column shows the results with household forecasts of inflation as the dependent variable and the right column shows the difference between the two. The first two parameters are taken from the estimations in the baseline case without any additional variables. Each of the other six parameters are taken from the estimations in the case in which the corresponding variable is v_t^i.

The response to the baseline variables is similar for realised and forecast inflation with no significant or important differences. In both cases inflation has a sensible average value[22] and a positive but low auto-correlation. Therefore there do not appear to be important conditional biases in the responses to the baseline information. The response to the three factors is much more interesting. Realised inflation reacts significantly and positively to realised inflation, suggesting that broad price rises are more sustained than narrow price increases. However forecast inflation reacts far less strongly to broad inflation, so there is a significant difference between the two, implying a large conditional bias. Comparing these coefficients to the equivalent estimated optimal coefficients from Table 1 also suggests that this bias in household forecasts is very sensible, as it is well within the band of the optimal shrunken

[22]Note that the average value is not just equal to the constant.

coefficients.

The response of realised inflation to the business cycle and financial cycle factors is positive. However the household forecasts barely respond to the business cycle factor and actually respond negatively to the financial cycle factor. There are therefore meaningful differences in the reactions to both factors, although only the financial cycle difference in statistically significant, implying conditional biases in the household forecasts. The lack of response to business cycle information is completely consistent with the optimal feasible expectation estimates in Table 1. However the response to financial cycle information is actually overly negative, suggesting that it may arise from a behavioural mistake[23], that leads to an even greater reduction in the coefficient than that required for optimal feasible expectations.

The response of realised inflation to exchange rates is clearly negative whereas household forecasts barely respond to exchange rates, suggesting a conditional bias although the difference between the two is just short of statistical significance. This conditional bias is completely in line with the conditional bias in optimal feasible expectations, as complete shrinkage is optimal in this case. Both wages and narrative federal funds futures markets monetary shocks only have very small associations with future inflation. They also have small associations with household forecasts, so there are no large conditional biases, although it is hard to comment on whether there are any conditional biases as the scale of their associations is too small to statistically detect biases with any confidence. Both very small coefficients are close to the equivalent coefficients in optimal feasible expectations, so are consistent with optimal feasible expectations.

These results suggest that the conditional biases in household inflation forecasts are very similar to the conditional biases in optimal feasible inflation expectations. In fact the only conditional bias that appears to be meaningfully different is that with respect to the financial cycle factor. Therefore many of the important nominal rigidities in the response of actual household inflation expectations to shocks will be the same as those in the response of optimal feasible inflation expectations.

[23] It could also be a small sample effect driven by mistaken expectations around the global financial crisis as this event was so important for this variable.

Table 2: Conditional biases in household inflation forecasts

	π^r_{t+h}	π^f_{t+h}	Difference
γ_0	1.91*	2.30*	0.39
	(1.32 to 2.46)	(2.08 to 2.54)	(-0.26 to 1.12)
γ_1	0.27*	0.29*	0.02
	(0.06 to 0.50)	(0.21 to 0.37)	(-0.26 to 0.28)
α^{inf}	2.08*	0.43	-1.64*
	(1.28 to 2.80)	(-0.07 to 0.80)	(-2.65 to -0.81)
α^{bc}	0.24	-0.04	-0.28
	(-0.07 to 0.50)	(-0.18 to 0.06)	(-0.61 to 0.07)
α^{fc}	0.30	-0.14*	-0.44*
	(-0.07 to 0.69)	(-0.23 to -0.07)	(-0.89 to -0.02)
α^{er}	-0.18	0.03	0.21
	(-0.44 to 0.03)	(-0.05 to 0.11)	(-0.03 to 0.51)
α^{w}	0.07	-0.01	-0.08
	(-0.26 to 0.35)	(-0.12 to 0.12)	(-0.41 to 0.34)
α^{nms}	0.03	0.04	0.02
	(-0.11 to 0.21)	(-0.02 to 0.09)	(-0.20 to 0.18)

Notes: Column 1 shows the OLS estimates of Equation 7 with realised inflation as the dependent variable, Column 2 shows the OLS estimates of Equation 8 with household inflation forecasts as the dependent variable and Column 3 shows the difference between the two coefficients. 90% block bootstrapped confidence intervals are in brackets and * = statistically significant at the 10% level.

For instance the response of household inflation expectations to a contractionary monetary shock is likely to be similar to the response of optimal feasible expectations discussed at the end of Section 4. Indeed, it may actually be more rigid as a result of household inflation expectations possibly reacting positively to any decline in financial cycle indicators.

Household expectations are not consistent with being entirely formed by optimal feasible expectations however, as they are clearly beaten in forecast performance by a feasible alternative. I show this by comparing the forecast performance of the forecasts produced by estimating my baseline auto-regressive forecasts on each of the training sets of data used in Section 3 with the equivalent household forecasts made at the end of each training set. The results show that the mean absolute forecast error of the household forecasts is 132% of that of the auto-regressive forecasts and the mean square forecast error of the household forecasts is 160% of that of the auto-regressive forecasts. This seems sensible as Figure 1 suggests that the household forecasts sometimes deviate persistently from realised inflation for years at a time. Therefore both optimal feasible inflation expectations and behavioural errors that reduce forecast performance seem to be important in explaining household inflation forecasts.

6 Conclusion

Inflation expectations and forecasts have a particular importance in economics and finance, as they affect the degree of nominal rigidities to aggregate demand shocks and so affect the size of their real impacts. The existing empirical literature has suggested that inflation expectations contain conditional biases with respect to publicly available macroeconomic information, causing nominal rigidities. However this is usually justified by behavioural factors, such as limited attention or imperfect cognitive abilities. While I do not deny the importance of these factors, in this paper I primarily study whether rational expectations are the optimal feasible expectations for agents, i.e. are they the expectations that are predicted to minimise a measure of forecast error out of the set of expectations that are feasible for agents to use.

I show that, with data samples of realistic length, agents will have to introduce conditional biases into their forecast in the vast majority of macroeconomic settings, due to the limited number of relevant monthly observations available relative to variables that can affect how prices are set. They can do this by shrinking their forecasts towards those given by a simpler specifications. However even if an agent had included sufficient shrinkage to make estimation feasible, it may well still be worth including additional shrinkage, as this may reduce the conditional forecast variance sufficiently to outweigh the increased conditional biases. Macroeconomic theory is unlikely to typically help to set the simpler specifications used in expectations formation, as its predictions are usually conditioned on state variables such as macroeconomic shocks and output gaps that are not well observed in real time. The importance placed on parsimony by the empirical forecasting literature and the degree to which auto-regressive benchmarks are hard to substantially beat in forecasting horse races, despite the no doubt numerous effects of many macroeconomic variables on each other, suggests that the extent of this optimal shrinkage and the conditional biases it causes could be very large in most applications.

Therefore rational expectations do not typically appear to be feasible for agents to learn from data and the optimal feasible expectations are likely to be very different to rational expectations in the vast majority of macroeconomic settings. As a result I suggest optimal feasible expectations as a new class of expectations. This suggests that optimal feasible expectations, with additional behavioural errors in some cases, could explain forecast formation across economics and finance.

The precise size of the conditional biases in optimal feasible expectations in any particular setting is partly an empirical question. I therefore empirically examine the size of the conditional biases in estimated optimal feasible expectations of US inflation. I do this by starting with a sensible auto-regressive benchmark and then consider the degree of shrinkage that it is optimal to apply to information on six important macroeconomic and financial variables using pseudo out of sample forecast performance. The variables I consider are a combined business cycle indicator, a combined financial cycle indicator, a combined indicator of broad inflation, trade-weighted exchange rates, hourly wages and narrative monetary shocks. I find that it is optimal to apply a very high degree of shrinkage to the most of these variables. Indeed it is optimal to apply absolute shrinkage to the business cycle indicator, the financial cycle indicator, exchange rates and wages, so this information is not included in the forecasts produced. It is also optimal to apply partial shrinkage to the broad inflation series but none to the narrative monetary shocks, although the shocks only have small associations with future inflation and the sign on these is not that implied by theory for true monetary shocks. However some of the macroeconomic variables having economically and statistically significant associations with future inflation, so the results imply that there are large conditional biases in optimal feasible expectations. Optimal feasible inflation expectations therefore appear to be very far from rational inflation expectations and are likely to contain large conditional biases that cause significant nominal rigidities in the reactions to macroeconomic shocks.

I also examine the conditional biases in surveys of US households' inflation forecasts and compare them to the conditional biases in estimated optimal feasible expectations. I find that household forecasts have a much smaller association with the broad inflation index than future realised inflation does and barely react to most of the other variables, although they do have a statistically significant but incorrect association with the financial cycle indicator. Therefore their conditional biases appear to be very similar to the conditional biases in optimal feasible expectations, except with regards to financial cycles. As a result the household inflation expectations are likely to produce similar, or even greater, levels of nominal rigidities in response to macroeconomic shocks than optimal feasible expectations would. However I also confirm that household expectations are not consistent with being entirely formed by optimal feasible expectations by showing that they are clearly beaten in a pseudo out of sample forecasting horse race by a feasible alternative: my auto-regressive benchmark. This may be caused by several persistent and seemingly unjustified deviations of household inflation forecasts from realised inflation that can last for years at a time as well as a mistaken reaction to financial cycle information. Therefore household forecasts of inflation expectations appear to be well explained by a combination of optimal feasible inflation expectations and behavioural mistakes that reduce forecast performance.

Optimal feasible expectations are therefore likely to cause important nominal rigidities, but they also have important implications across economics and finance. For instance, they may imply that agents forecasts of future asset returns are conditionally biased towards the long-term average return of similar assets. Hence they provide one reason why agents fail to adjust asset demand to remove small associations between current variables and future risk-adjusted returns and so could fuel asset price bubbles. They may also imply that agents forecasts of the probabilities of being the pivotal voter between each combination of two plausible candidates in an election are biased towards a single probability in electoral systems that make these probabilities hard to predict, limiting the applicability of Arrow's impossibility theorem. Exploring these implications is an important area for future research.

Taken together, these results suggest that inflation forecasts, as well as forecasts across economics and finance, are formed by a combination of optimal feasible expectations and, in some cases, additional behavioural errors. Therefore macroeconomic and financial models based on rational expectations may be fundamentally misspecified and need to be dramatically changed to incorporate realistic forecasting behaviour.

Appendices

A Variables used to construct factors

The following tables contain a complete list of the variables used in this paper and the transformations applied to them. APC stands for annual percentage change and the sources and transformations are discussed in Section 3.

Table 3: List of dependent, non-factor and price factor variables

	Transformation	Source
CPI: Index	APC	Fred MD
Annual inflation expectations	None	Michigan Consumer Survey
Trade-weighted exchange rate index	APC	Fred MD
Average hourly earnings	APC	Fred MD
Narrative monetary shocks	None	Ramey (2016)
PPI: Finished Goods	APC	Fred MD
PPI: Finished Consumer Goods	APC	Fred MD
PPI: Intermediate Materials	APC	Fred MD
PPI: Crude Materials	APC	Fred MD
Crude Oil, spliced WTI and Cushing	APC	Fred MD
PPI: Metals and metal products	APC	Fred MD
CPI: Apparel	APC	Fred MD
CPI: Transportation	APC	Fred MD
CPI: Medical Care	APC	Fred MD
CPI: Commodities	APC	Fred MD
CPI: Durables	APC	Fred MD
CPI: Services	APC	Fred MD
CPI: All Items Less Food	APC	Fred MD
CPI: All items less shelter	APC	Fred MD
CPI: All items less medical care	APC	Fred MD
PCE: Chain index	APC	Fred MD
PCE: Durable goods	APC	Fred MD
PCE: Nondurable goods	APC	Fred MD
PCE: Services	APC	Fred MD
CPI: Food at home	APC	BLS
CPI: Food away from home	APC	BLS
CPI: Rent of primary residence	APC	BLS
CPI: Fuel and utilities	APC	BLS
CPI: New and used motor vehicles	APC	BLS
CPI: Motor fuel	APC	BLS
CPI: Medical care services	APC	BLS
CPI: Other goods and services	APC	BLS
PCE: Excluding food and energy	APC	BEA
PCE: Energy goods and services	APC	BEA
PCE: Food	APC	BEA
Sticky price index	APC	Atlanta Fed
Sticky price index less food and energy	APC	Atlanta Fed
Sticky price index less shelter	APC	Atlanta Fed
Flexible price index	APC	Atlanta Fed
Flexible price index less food and energy	APC	Atlanta Fed

Table 4: List of business cycle factor variables

	Transformation	Source
Real Personal Income	APC	Fred MD
Real personal income ex transfer receipts	APC	Fred MD
IP Index	APC	Fred MD
IP: Final Products and Nonindustrial Supplies	APC	Fred MD
IP: Final Products	APC	Fred MD
IP: Consumer Goods	APC	Fred MD
IP: Durable Consumer Goods	APC	Fred MD
IP: Nondurable Consumer Goods	APC	Fred MD
IP: Business Equipment	APC	Fred MD
IP: Materials	APC	Fred MD
IP: Durable Materials	APC	Fred MD
IP: Nondurable Materials	APC	Fred MD
IP: Manufacturing	APC	Fred MD
IP: Residential Utilities	APC	Fred MD
IP: Fuels	APC	Fred MD
Civilian Labor Force	APC	Fred MD
Civilian Employment	APC	Fred MD
Civilians Unemployed - <5 weeks	APC	Fred MD
Civilians Unemployed - 5-14 weeks	APC	Fred MD
Civilians Unemployed - >14 weeks	APC	Fred MD
Civilians Unemployed - 15-26 weeks	APC	Fred MD
Civilians Unemployed - >27 weeks	APC	Fred MD
Initial Claims	APC	Fred MD
All Employees: Total nonfarm	APC	Fred MD
All Employees: Goods-Producing	APC	Fred MD
All Employees: Mining	APC	Fred MD
All Employees: Construction	APC	Fred MD
All Employees: Manufacturing	APC	Fred MD
All Employees: Durable goods	APC	Fred MD
All Employees: Nondurable goods	APC	Fred MD
All Employees: Service-Providing	APC	Fred MD
All Employees: Transport and others	APC	Fred MD
All Employees: Wholesale Trade	APC	Fred MD
All Employees: Retail Trade	APC	Fred MD
All Employees: Financial Activities	APC	Fred MD
All Employees: Government	APC	Fred MD

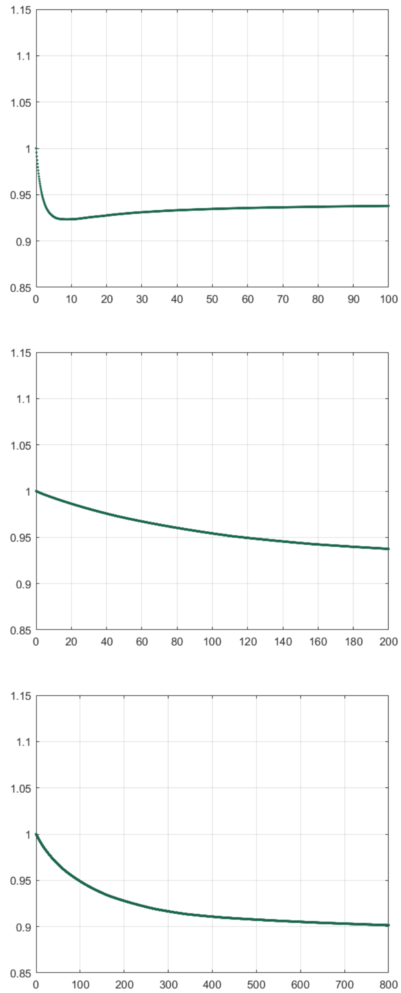
Table 5: List of financial cycle factor variables

	Transformation	Source
Commercial and Industrial Loans	APC	Fred MD
Real Estate Loans at All Commercial Banks	APC	Fred MD
Total Nonrevolving Credit	APC	Fred MD
Nonrevolving consumer credit to Personal Income	APC	Fred MD
Consumer Motor Vehicle Loans Outstanding	APC	Fred MD
Total Consumer Loans and Leases Outstanding	APC	Fred MD
Securities in Bank Credit at All Commercial Banks	APC	Fred MD
Total Consumer Credit Owned and Securitized	APC	Fed Board
Total Revolving Credit Owned and Securitized	APC	Fed Board
House prices: National index	APC	Freddie Mac
House prices: New York index	APC	Freddie Mac
House prices: Los Angeles index	APC	Freddie Mac
House prices: Chicago index	APC	Freddie Mac
House prices: Dallas index	APC	Freddie Mac
House prices: Houston index	APC	Freddie Mac
House prices: Washington index	APC	Freddie Mac
House prices: Miami index	APC	Freddie Mac
House prices: Philadelphia index	APC	Freddie Mac
House prices: Atlanta index	APC	Freddie Mac
House prices: Boston index	APC	Freddie Mac
House prices: Phoenix index	APC	Freddie Mac
House prices: San Francisco index	APC	Freddie Mac
House prices: Riverside index	APC	Freddie Mac
House prices: Detroit index	APC	Freddie Mac
House prices: Seattle index	APC	Freddie Mac
House prices: Minneapolis index	APC	Freddie Mac
House prices: San Diego index	APC	Freddie Mac
House prices: Tampa index	APC	Freddie Mac
House prices: Denver index	APC	Freddie Mac
House prices: St Louis index	APC	Freddie Mac
House prices: Baltimore index	APC	Freddie Mac
House prices: Orlando index	APC	Freddie Mac
House prices: Charlotte index	APC	Freddie Mac
House prices: San Antonio index	APC	Freddie Mac
House prices: Portland index	APC	Freddie Mac
House prices: Sacramento index	APC	Freddie Mac
House prices: Pittsburgh index	APC	Freddie Mac
House prices: Las Vegas index	APC	Freddie Mac
House prices: Cincinnati index	APC	Freddie Mac
House prices: Austin index	APC	Freddie Mac
House prices: Kansas index	APC	Freddie Mac

B Training set size robustness

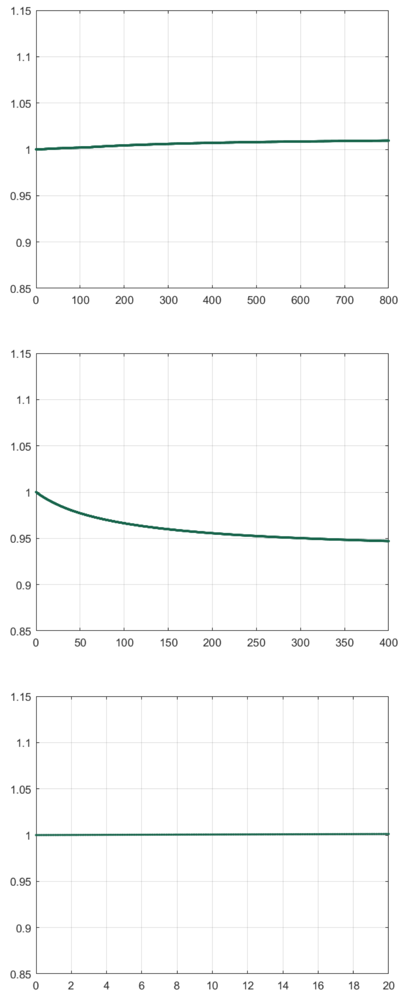
This section of the Appendix contains the robustness checks from using different proportions of the sample as the training sample. In particular, it repeats analysis from Section 4 but takes either 60% or 80% of the full sample as the training sample, instead of the 70% in the baseline analysis.

Figure 6: Forecast error with shrunken inflation factor, business cycle factor and financial cycle factor



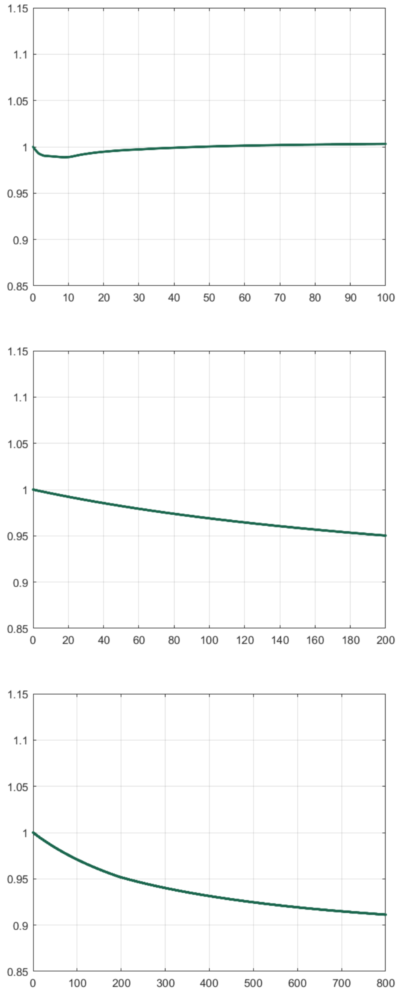
Notes: This figure repeats the graphs in Figure 4 but with estimates based on training sample of 60% of the total dataset.

Figure 7: Forecast error with shrunken exchange rates, wages and narrative federal funds market monetary shocks



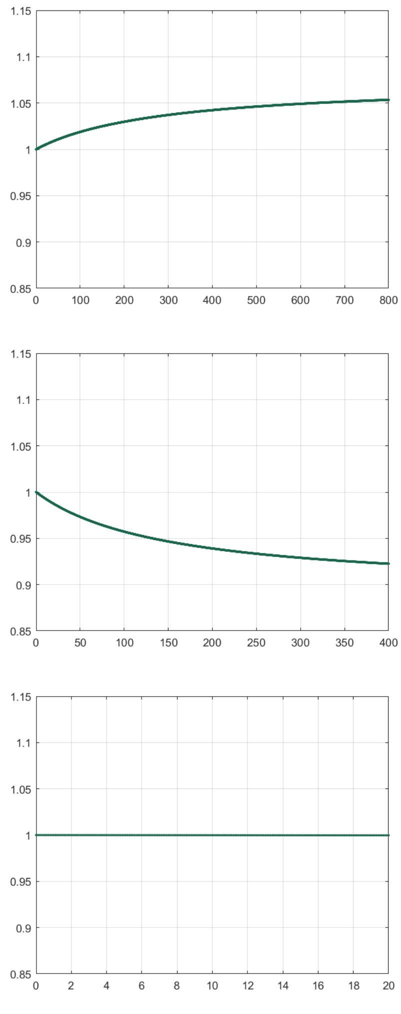
Notes: This figure repeats the graphs in Figure 5 but with estimates based on training sample of 60% of the total dataset.

Figure 8: Forecast error with shrunken inflation factor, business cycle factor and financial cycle factor



Notes: This figure repeats the graphs in Figure 4 but with estimates based on training sample of 80% of the total dataset.

Figure 9: Forecast error with shrunken exchange rates, wages and narrative federal funds market monetary shocks

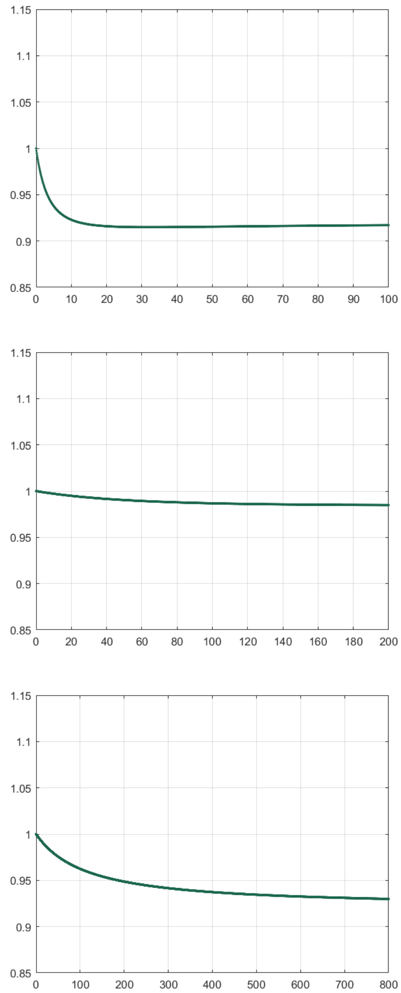


Notes: This figure repeats the graphs in Figure 5 but with estimates based on training sample of 80% of the total dataset.

C Forecast performance measure robustness

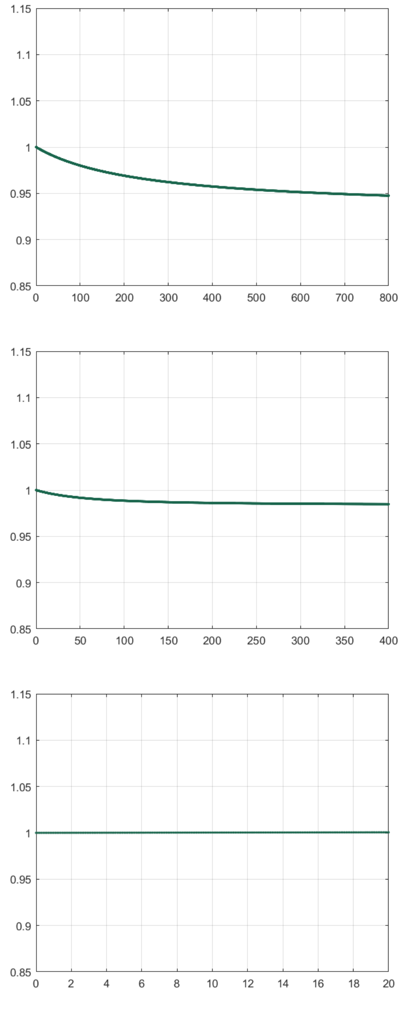
This section of the Appendix contains the robustness checks from using a different measure of forecast performance. In particular, it repeats analysis from Section 4 but uses the mean square forecast error instead of the mean absolute forecast error as the measure of out of sample forecast performance.

Figure 10: Forecast error with shrunken inflation factor, business cycle factor and financial cycle factor



Notes: This figure repeats the graphs in Figure 4 but uses the mean square forecast error instead of the mean absolute forecast error.

Figure 11: Forecast error with shrunken exchange rates, wages and narrative federal funds market monetary shocks



Notes: This figure repeats the graphs in Figure 5 but uses the mean square forecast error instead of the mean absolute forecast error.

References

Ahmed, N. K., Atiya, A. F., El Gayar, N., and El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 29(5-6):594–621.

Akerlof, G. A. (2002). Behavioral Macroeconomics and Macroeconomic Behavior. *American Economic Review*, 92(3):411–433.

Bai, J. and Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, 74(4):1133–1150.

Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317.

Berge, T. J. (2018). Understanding survey-based inflation expectations. *International Journal of Forecasting*, 34(4):788–801.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120(1):387–422.

Breitung, J. and Eickmeier, S. (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, 163(1):71–84.

Chung, H. and Xiao, W. (2013). Cognitive Consistency, Signal Extraction, and Macroeconomic Persistence. *SUNY Binghampton Working Paper*.

Coibion, O. and Gorodnichenko, Y. (2012a). What Can Survey Forecasts Tell Us About Informational Rigidities? *Journal of Political Economy*, 120(1):116–159.

Coibion, O. and Gorodnichenko, Y. (2012b). Why Are Target Interest Rate Changes so Persistent. *American Economic Journal: Macroeconomics*, 4(4):126–162.

Coibion, O. and Gorodnichenko, Y. (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.

Coibion, O. and Gorodnichenko, Y. (2015b). Is the phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232.

Coibion, O., Gorodnichenko, Y., and Kamdar, R. (2018). The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature*, 56(4):1447–91.

De Mol, C., Giannone, D., and Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2):318–328.

Del Giovane, P., Fabiani, S., and Sabbatini, R. (2008). What's behind inflation perceptions? A survey-based analysis of Italian consumers. *Bank of Italy Working Papers*, No. 655.

Diebold, F. X. (2017). *Forecasting*. University of Pennsylvania, Philadelphia.

Evans, G. W. and Honkapohja, S. (2012). *Learning and expectations in macroeconomics*. Princeton University Press.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In *Handbook of Economic Forecasting*. Elsevier.

Friedman, M. (1977). Nobel Lecture: Inflation and Unemployment. *Journal of Political Economy*, 85(3):451–472.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Giacomini, R. (2015). Economic theory and forecasting: lessons from the literature. *Econometrics Journal*, 18(2):C22–C41.

Gordon, R. J. (1982). Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment. *National Bureau of Economic Research Working Paper Series*, No. 708.

Granger, C. and Newbold, P. (1986). *Forecasting Economic Time Series*. Academic Press.

Havranek, T. and Rusnak, M. (2013). Transmission lags of monetary policy: A meta-analysis. *International Journal of Central Banking*, 9(4):37–76.

Hommes, C., Mavromatis, K., Ozden, T., and Zhu, M. (2019). Behavioral Learning Equilibria in the New Keynesian Model. *De Nederlandsche Bank Working Paper Series*, No. 654.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. Palgrave Macmillan.

Kim, H. H. and Swanson, N. R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34(2):339–354.

Koop, G. (2017). Bayesian methods for empirical macroeconomics with big data. *Review of Economic Analysis*, 9(1):33–56.

Lovell, M. (1986). Tests of the Rational Expectations Hypothesis. *The American Economic Review*, 76(1):110–124.

Lucas, R. E. (1996). Nobel Lecture: Monetary Neutrality. *Journal of Political Economy*, 104(4):661–682.

Medeiros, M. C., Vasconcelos, G., Veiga, A., and Zilberman, E. (2020). Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods. *Journal of Business & Economic Statistics (forthcoming)*.

Miranda-Agrippino, S. (2016). Unsurprising shocks: information, premia, and the monetary transmission. *Bank of England Staff Working Papers*, No. 626.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, pages 313–315.

Orphanides, A. and Williams, J. C. (2007). Robust monetary policy with imperfect knowledge. *Journal of Monetary Economics*, 54(5):1406–1435.

Pesaran, M. H. and Weale, M. (2006). Survey Expectations. In *Handbook of Economic Forecasting*, pages 715–776. North Holland.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. In *Handbook of Macroeconomics*, volume 2, pages 71–162. Elsevier.

Sheffrin, S. M. (1996). *Rational Expectations.* Cambridge University Press.

Stock, J. H. and Watson, M. W. (1996). Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business & Economic Statistics*, 14(1):11–30.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147–162.

Stock, J. H. and Watson, M. W. (2008). Phillips Curve Inflation Forecasts. *National Bureau of Economic Research Working Papers*, No. w14322.

Stock, J. H. and Watson, M. W. (2016). Core inflation and trend inflation. *Review of Economics and Statistics*, 98(4):770–784.