Cross-Country Inflation Expectations
Evidence of Heterogeneous & Synchronized 'Mistakes'

Karen Smith[1]

November 8, 2022

Latest Version: Please click here

Abstract

This paper presents a cross-country horizon analysis of departures from full-information rational expectations (FIRE) among professional forecasters from 18 OECD countries. Using four well-known tests of FIRE, I find significant heterogeneity in the cross-section of results. My findings contradict the existing literature which has often generalized the direction and magnitude of departures from FIRE. I show that coexisting with these cross-country variations is a common factor in forecast errors – or a cycle of synchronized 'mistakes'. I argue that this factor may have contributed to the one-size-fits-all approach to the characterization of the expectation formation process.

[1]Brandeis University, International Business School. Email: klsmith@brandeis.edu

1 Introduction

The main contribution of this paper is to present a comparative analysis of cross-country departures from full-information rational expectations (FIRE) and to demonstrate that despite the heterogeneity in these results, there is significant co-movement in the inflation expectations of professional forecasters. Using survey data from forecasters in 18 advanced economies, I conduct a cross-sectional analysis of their predictions by forecast horizon. While it is established that survey forecasts of agent expectations violate the assumptions of full-information rational expectations, what is less known is the extent to which these violations vary in the cross-section. Evidence from this analysis suggests that generalizations about the outcome of well-known tests of rational expectations are not ubiquitous. I find that differences in the magnitude and direction of forecaster 'mistakes' often contradict the existing literature. The implication is that there is no one-size-fits-all approach to how we account for departures from full-information rational expectations and more importantly, how we model the expectation formation process.

I begin this analysis by conducting four widely-used tests of FIRE on monthly survey forecasts of the average annual inflation rate from Consensus Economics – an international forecasting firm – over the period January 1990 to December 2020. The unique structure of these forecasts facilitates a multi-horizon analysis of the results across countries. I confirm the occurrence of cross-country violations of FIRE by professional forecasters. This finding, however, is accompanied by significant variation in the outcomes of the tests. In one of my most striking results I show that for some countries such as the UK, Italy, Netherlands, Portugal, and Sweden, consensus forecasts concurrently exhibit under- as well as over-reaction at different forecast horizons.

Within the existing literature, however, empirical studies have largely revealed that the aggregate expectations of agents are characterized by under-reaction, which is often invariant to the forecast horizon (see Coibion and Gorodnichenko (2015) [13], Bordalo et *al.* (2020) [7]). This has led to a general acceptance of the view that aggregate expectations are characterized by inertia. My findings, therefore, represent evidence in the opposite direction of the conclusions of existing research. I challenge this widely-accepted view and discuss how earlier conclusions about the expected outcome of more recent tests of FIRE likely reflected country-specific factors.

I complement this analysis with the introduction of a newly digitized real-time cross-country inflation data set. It is well known that over time, macroeconomic data are subject to revisions, methodological changes, and even release lags which have the potential to materially affect the results of the analysis upon which such data rests. Real-time data are data that have not been subject to such adjustments.[2] In the context of this paper, these data are used as the benchmark to evaluate historical inflation predictions in line with the information that would have been available to the forecaster at the time the forecast was made.

[2] See Croushore and Stark (2001) [14] for a discussion on the advantages of using real-time data

As one of the primary sources of international real-time data, the OECD maintains a database of 21 key economic variables from 1999 onward. To conduct this study, I augment the availability of international real-time CPI data by digitizing 10 years of monthly hard copy records of the OECD's Main Economic Indicators (MEI) publication from 1989 to 1999, complete with the accompanying vintages.[3] These records were gathered from various public sources and as an outcome of this study, will be made freely available to anyone interested in international macroeconomic research.

As a further contribution, this paper provides a basis for comparatively evaluating the consistency of theoretical tests of full-information rational expectations. My initial finding is that the heterogeneity in results across countries and forecast horizons is also reflected in the disparate messages conveyed by the results of each test. For example, while one test may predominantly convey over-prediction of inflation relative to its realized value, another may broadly convey under-reaction by forecasters. Upon closer examination, however, I discuss how these seemingly contrasting findings may be connected and even reconciled. I hypothesize that cross-country departures from FIRE may in part be driven by commonalities and that the existence of these commonalities may account for many of the generalizations made about test results within the literature.

To test this hypothesis I introduce a Bayesian Dynamic Factor Model (BDFM) of 'synchronized mistakes'. This model is distinguished from the commonly used static factor model as it incorporates the inter-temporal cross-correlations among the observables. Following Stock and Watson (1989, 1993) [46], [44] and Bai and Wang (2015) [5], I introduce two key restrictions which identify the model. I then use the Gibbs sampler, a Monte Carlo Markov chain (MCMC) procedure, to estimate the joint posterior distribution of model parameters.

The results point, overwhelmingly, to the existence of a common latent factor driving cross-country forecast errors over various forecast horizons, throughout this study. This cycle of forecaster 'mistakes' emphasizes the importance of international factors in national macroeconomic developments. The results show that the cross-country factor was particularly important to domestic forecast errors during economic shocks such as the recession of the early 2000s and the Global Financial Crisis (GFC). Relatedly, the results reveal an apparent oscillation of the cross-country factor over time, characterized by periods of over- and under-prediction in forecasts. I find that in periods leading up to the economic shocks, the common factor reflects positive forecast errors – under-prediction – and that this result reverses, albeit asymmetrically, during the subsequent period of economic decline.

Similar to the approach used in conducting the tests of FIRE, the dynamic factor analysis also facilitates key comparisons across countries. I extend the analysis by conducting a variance decomposition of the cross-country latent factor in forecast errors and find that on average approximately 43% and 40% of the volatility in domestic forecast errors at the respective one- and two-year-ahead forecast horizons, are driven by variation in the cross-country factor. The

[3] It should be noted here that the Federal Reserve Bank of Dallas also maintains a Real-Time Historical Database for the OECD from 1962 to 1998 which can be merged with the OECD's data which begins from 1999. However, these data are provided only quarterly, whilst the OECD's data are monthly.

inverse of this result is that approximately 57% and 60% of the aggregate variation in domestic forecast errors were driven by home-country inflation volatility throughout this study.

To more closely examine the heterogeneous effects of the cross-country factor, I turn to the factor loadings of the model. These loadings communicate the relative importance of domestic factors in expectational errors. Based on my analysis, I find significant variation across countries. At the one-year-ahead forecast horizon, for example, Norway's loading on the cross-country factor is only 0.36, whereas, France places a weight of 2.24 on the factor. This disparity between countries occurs at every forecast horizon and reinforces the initial findings of heterogeneity in departures from full-information rational expectations. In other words, forecaster 'mistakes' are also distinctly driven by factors specific to domestic conditions.

In summary, this paper presents an analysis of the cross-country departures from full-information rational expectations. Several tests of rational expectations reveal significant variation in the extent to which cross-country expectations depart from FIRE across forecast horizons. I find, however, that coexisting with these heterogeneous results are several commonalities. I discuss how these contrasting results can be reconciled. Through my application of dynamic factor analysis, I provide evidence of the existence of a common factor in forecast errors and argue that this factor likely accounts for commonalities observed in deviations from FIRE and the concomitant generalizations in the literature. I also show, that country-specific factors also feature prominently in the common factor.

Together, these findings point to the need for a more global approach to understanding the expectation formation process. The broader implication of this study is that given the increasingly interconnected global environment, and the well-established endogeneity of inflation expectations, decision-makers must consider, more strongly, the impact of international factors on domestic macroeconomic policy decisions.

1.1 Related Literature

This paper is related to three main strands of literature. Firstly, this paper challenges the underlying assumptions of the longstanding rational expectations revolution put forward by Muth (1961) [36], Lucas (1976) [28], and Lucas and Sargent (1981) [24]. The full-information rational expectations hypothesis holds that agents know and understand the true model that governs the economy, they can observe the associated macroeconomic shocks, and form expectations based on these observations. Since their beliefs are fully in sync with the model which specifies the evolution of the economy, their expectations are statistically optimized forecasts. The implication of these assumptions is that forecast errors are unpredictable from any information available at time t or earlier. This conclusion forms the basis for several tests of the null hypothesis of full-information rational expectations.

There is a substantial body of research that point to the rejection of FIRE based on empirical findings from survey data. For example, Jonung and Laidler (1988) [23] in a Swedish survey of households' contemporaneous inflation expectations, find that forecast errors of survey participants are serially correlated. In a more broad-based approach, Mankiw, Reis, and Wolfers (2003) [32] analyze over 50 years of data on inflation expectations across various well-known surveys in the US.[4] Using different tests of rational expectations, the authors show that there is no single survey that fully satisfies the assumptions of FIRE.

In tests of the sub-samples of the population, FIRE is also rejected. Souleles (2004) [43] finds that forecast errors from the Michigan Index of Consumer Sentiment are correlated with consumer demographics. In the more recent literature, Bordalo et *al.* (2020) [7] apply a more comprehensive approach by testing departures from FIRE across 22 macroeconomic and financial variables from the Survey of Professional Forecasters (SPF) and the Blue Chip Financial Forecasts. The authors confirm the predictability of forecast errors from forecast revisions – a clear violation of FIRE.

Empirically, this paper also contributes to this growing literature by providing a relatively standardized analysis of departures from FIRE across countries, amongst the same type of forecasting agents – professional forecasters. To my knowledge, there are no other papers that adopt such a broad-based approach to studying the expectation formation process. The objective is to provide a baseline for future evaluation of FIRE deviations in an international context.

There are several econometric tests of the null hypothesis of full-information rational expectations that have emerged over the years. As noted by Sheffrin (1996) [41], all tests of FIRE are essentially tests of the underlying properties of conditional expectations and can be classified accordingly.[5] To facilitate clearer comparisons, I classify these tests into two broad categories – tests of unbiasedness and tests of efficiency.

Unbiasedness tests such as the Mincer-Zarnowitz Test (1989) [33] propose that on average, agents should forecast the variable precisely or with zero prediction error. On the other hand, efficiency tests evaluate the agent's ability to incorporate all relevant information into the forecast. These tests generally take the form of regressing the forecast or the forecast error on the variables in the information set of the forecaster. By construction, they follow from the orthogonality property of forecast errors which holds that these errors should be uncorrelated with any information available to agents at time t or earlier. One such test is that put forward by Coibion and Gorodnichenko (2015) [13] in which the authors regress the h-period ahead forecast error against the time t forecast revision. A test of the auto-correlation of forecast errors also falls within this classification.

I add to this extensive literature the dimension of cross-sectional horizon analysis of the tests of departures from FIRE by holding constant the period over which the tests are conducted and varying the forecast horizon. While this approach has traditionally been used in the literature to gauge disagreement amongst forecasters (Patton and Timmermann (2010) [39]), the application in this paper differs from the traditional horizon analysis because as it reveals more dynamically how forecaster deviations from FIRE strengthen or attenuate over time and across

[4]These surveys include, the University of Michigan's Survey of Consumers, the Survey of Professional Consumers and Livingston Survey.

[5]Sheffrin (1996) [41] notes that in keeping with the properties of conditional expectations, tests of rational expectations can be classified into four main categories: (1) Unbiasedness, (2) Efficiency, (3) Forecast Error Predictability, and (4) Consistency.

countries. This provides the unique opportunity to gain further insight into the applicability of traditional models used to account for violations of FIRE.

This paper also relates to the literature which analyzes the importance of cross-country linkages in macroeconomic fluctuations. Similarities in business-cycle fluctuations, asset pricing, and inflation dynamics are just some of the areas where researchers have sought to establish the relationship between domestic macroeconomic developments and international conditions. For example, Kose et *al.* [27] provide evidence that a common world factor accounts for a significant source of the variation in aggregates in many domestic economies – leading to the conclusion of the existence of a global business cycle. Miranda-Agrippano and Rey (2020) [34] show that US monetary policy shocks induce a 'Global Financial Cycle' due to the existence of a cross-country factor that accounts for 20% of the variation in international asset prices.

In the context of inflation, it has long been argued that domestic inflation is driven by a global factor. Borio and Filardo (2007) [8] discuss how the integration of the global economy through globalization has contributed to inflation becoming more of a "globe-centric" than a "country-centric" phenomenon. The authors conclude that domestic inflation is more readily explained by the global output gap and that periods of low and stable inflation such as during the 1990s and 2000s coincide with a decline in the global economic slack.

Complementing these results are the findings of Ciccarelli and Mojon (2010) [12] which emphasize the role of the common factor. The authors find that this factor drives approximately 70% of the variation in domestic inflation of 22 advanced economies. This paper revisits the role of the cross-country factor in domestic inflation but from the perspective of inflation expectations.

Finally, this paper relates to the literature which advances the relevance of real-time international data. Among economists, there is consensus that the analysis of information known to agents at the time of economic decision-making leads to substantially different and more accurate conclusions compared to working with revised data (see Croushore and Stark (2001, 2003)[14][15], Orphanides (2001)[37], Koeing (2003)[26] and Molodtsova et *al.* (2008)[35]). The use of these data has gained prominence in areas such as economic forecasting, and monetary policy analysis.

While there has been significant progress in the availability of real-time data for the US, the availability of consistent and reliable international real-time macroeconomic data sets, has been especially limited.[6] To my knowledge, there are two main sources of international real-time data, both of which rely on historical data from the OECD's Main Economic Indicators (MEI), a monthly publication dating back to the OECD's inception in 1961. The first is the OECD's Revisions-Analysis data-set, which provides vintages of monthly real-time data from 1999 forward. Complementing this data-set is the Dallas Fed's real-time data compilation by Fernandez et *al.* (2011) [19]which contains quarterly vintages from 1962 to 1998 for 26 OECD

[6]For example, Croushore and Stark's (2001) Real-Time Data-set for Macroeconomists (RTDSM) administered by the Philadelphia Fed, contains data on the US economy from 1965 onward. The Federal Reserve Bank of St. Louis also maintains ALFRED which is an archive of real-time data from 1969 for the US and some advanced economies.

countries.

This paper adds to the existing literature on the relevance of cross-country real-time macroeconomic data in two important ways. Firstly, it presents one of the first empirical applications of international real-time data which facilitates comparative analysis of departures from full-information rational expectations. Secondly, this paper augments the availability of real-time CPI data for periods prior to 1999. In the section which follows I describe in greater detail, the manner in which this data-set is used.

2 Data

To conduct this empirical analysis I combine two main types of data on cross-country inflation - survey forecasts from professional forecasters and real-time CPI.

Forecast Data. My forecast data are sourced from Consensus Economics' international surveys of economic information. The data set contains forecasts of a range of macroeconomic and financial variables at both the consensus (aggregate) and individual forecaster levels for several countries. This study utilizes consensus forecasts from 18 advanced economies, including the G7 countries, from January 1990 to December 2020. Table (1) below lists the countries.

Table 1. Country List

USA	Canada	Netherlands
Japan	Austria	Norway
Germany	Belgium	Portugal
France	Denmark	Spain
UK	Finland	Sweden
Italy	Ireland	Switzerland

At the beginning of each month, forecasters from each country make predictions on the average annual percentage change in their respective consumer price indices. Forecasts are made for the current and subsequent calendar years. This data-set, therefore, provides 24 monthly current-year and next-year forecasts for each country throughout the study. In a similar approach to Patton and Timmermann (2010) [39], this paper utilizes the unique structure of these data to conduct a horizon analysis of the cross-section of forecasts.

The monthly forecast horizons used in this analysis are represented by $h = 1,...,24$ months, where, for example, forecast horizon $h = 12$ represents the forecast made in January of the current calendar year, for the current calendar year and $h = 24$ indicates the forecast of the average annual inflation rate made in January of the current calendar year for the subsequent year. Therefore, the time series of forecasts at each horizon utilizes 31 observations per country - one for each year from 1990 to 2020.

Real-Time Data. Throughout this analysis real-time data are used to measure ex-post inflation across countries. As discussed in the introduction of this paper, these data are sourced from the OECD's Main Economic Indicators (MEI) Original Release Data and Revisions Database and are available from February 1999 onward.[7] As an additional project and complementary to this analysis, I extended this data-set for the CPI variable for the period January 1989 to January 1999 for all 18 countries in this study. The data were sourced from hard copy publications of the OECD's Main Economic Indicators publications which were available from the public domain in the corresponding vintage.

Together, the forecast and real-time data are used to assess the historical performance of cross-country forecasts made at time t. Specifically, in the following section, I use these data to test for departures from full-information rational expectations which in many instances involves the construction of forecast errors. Therefore, the question of which vintage of the real-time data should be used to make the appropriate comparison to forecasts may arise. For this empirical application, I follow Coibion and Gorodnichenko (2015) [13] by comparing all forecasts to real-time data available one year after the period being forecasted over.

Summary Statistics. Figure (1) shows the consensus forecast for each country at selected forecast horizons whilst Table (2) provides additional summary statistics.

Figure 1. Consensus Forecasts of Average Annual Inflation by Country (%)

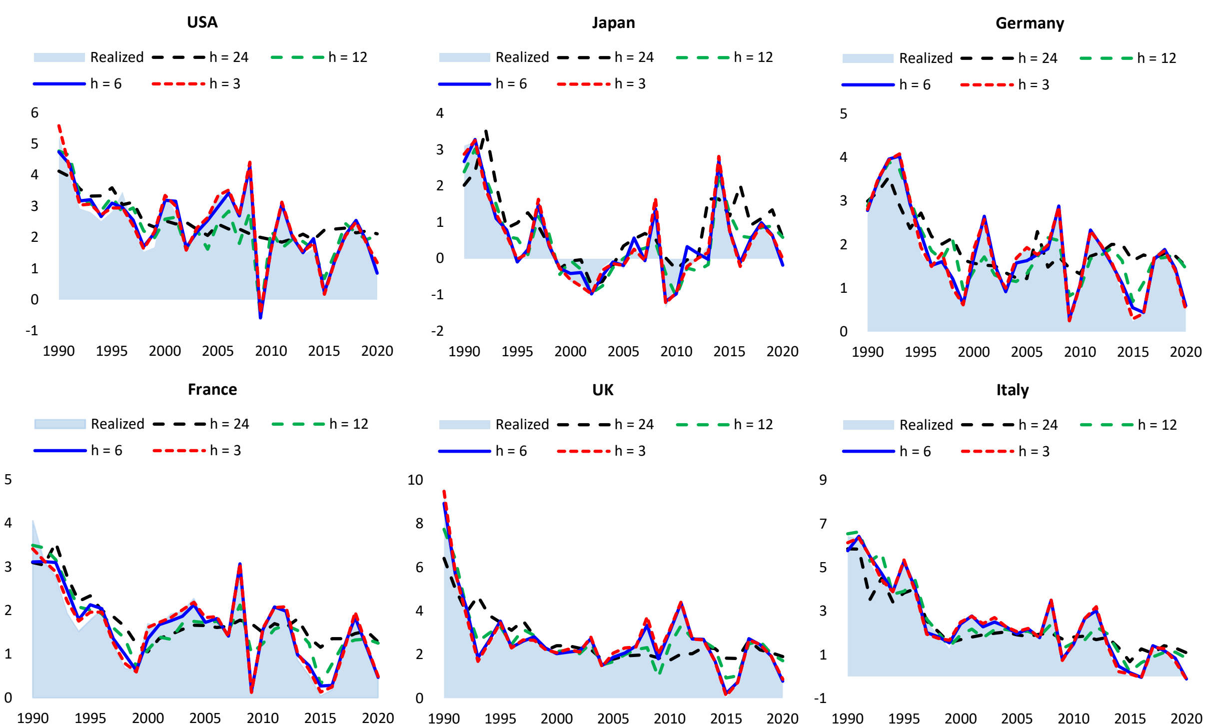
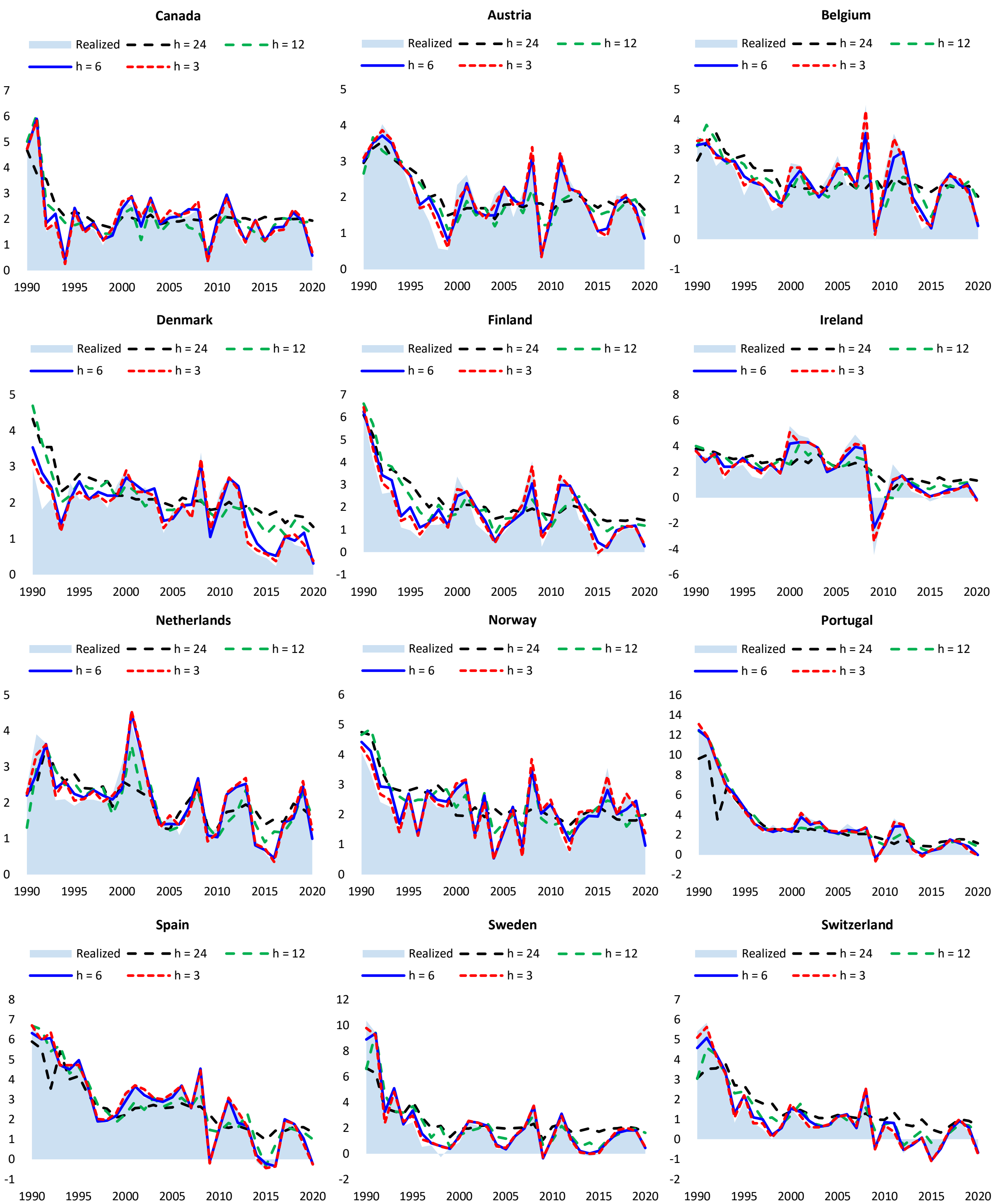


Figure 1. continued on the next page

[7]These data are available under 'General Statistics' at https://stats.oecd.org/

Figure 1. Consensus Forecasts of Average Annual Inflation by Country (%)



Note: The figure above shows annual consensus forecasts at forecast horizons, 24, 12, 6 and 3 months, along with realized inflation for each country in percentage points. All consensus forecasts sourced from Consensus Economics. All realized data are real-time sourced from the OECD's Main Economic Indicators Original Release Database and the author's compilation.

Table 2. Summary Statistics on Realized and Forecasted Inflation

	Standard Deviation					Root Mean Squared Error (RMSE)			
	Realized	Forecasts				Forecasts			
		h = 3	h = 6	h = 12	h = 24	h = 3	h = 6	h = 12	h = 24
USA	1.10	1.19	1.14	1.00	0.63	0.21	0.27	0.62	1.05
Japan	1.05	1.13	1.09	1.00	0.88	0.12	0.23	0.45	0.87
Germany	0.93	1.00	0.97	0.81	0.59	0.08	0.17	0.48	0.87
France	0.80	0.86	0.84	0.72	0.56	0.17	0.25	0.49	0.79
UK	1.51	1.66	1.58	1.38	1.11	0.10	0.19	0.63	1.12
Italy	1.69	1.76	1.73	1.69	1.36	0.18	0.22	0.61	0.89
Canada	1.06	1.10	1.09	1.02	0.60	0.14	0.23	0.61	0.93
Austria	0.81	0.90	0.84	0.64	0.53	0.25	0.36	0.64	0.84
Belgium	0.80	0.92	0.81	0.66	0.49	0.21	0.40	0.89	1.01
Denmark	0.76	0.80	0.80	0.74	0.62	0.21	0.35	0.77	0.95
Finland	1.32	1.39	1.32	1.31	1.10	0.31	0.52	1.03	1.44
Ireland	1.56	1.86	1.64	1.20	0.88	0.43	0.67	1.54	1.89
Netherlands	0.84	0.91	0.88	0.71	0.51	0.22	0.31	0.55	0.84
Norway	0.89	0.91	0.90	0.86	0.80	0.15	0.34	0.72	1.03
Portugal	3.05	3.20	3.13	3.10	2.48	0.27	0.43	0.87	1.24
Spain	1.80	1.93	1.86	1.69	1.32	0.17	0.26	0.75	1.12
Sweden	2.13	2.31	2.25	1.86	1.26	0.28	0.47	1.16	1.64
Switzerland	1.44	1.60	1.50	1.21	0.93	0.14	0.30	0.75	1.26

Note: The table shows the average standard deviation and Root Mean Squared Error (RMSE) calculations for realized and forecasted inflation for each country for the period 1990 to 2020 by forecast horizon.

These data on the cross-country consensus forecasts convey a consistent message. Figure (1) shows that for each country, forecasts made at the most distant forecast horizons, $h = 12$, 24, display less variation than the most recent predictions, $h = 3$, 6. Forecasters appear to be able to more accurately predict average annual inflation by as early as six months before the end of the year. From a macroeconomic perspective, many of the countries appear to have emerged from a period of higher inflation in the early 1990s to lower levels of inflation by the end of the period.

The summary statistics in Table (2) confirm these observations. For example, the standard deviation of forecasts is lowest at horizons $h = 12$, 24, and highest at horizons, $h = 3$, 6. In fact, by $h = 3$, professional forecasters appear to make predictions that almost exactly match the volatility in the realized values of inflation for their respective countries. These observations are consistent with the Root Mean Squared Error (RMSE) calculations for each country, which is significantly lower at the shortest forecast horizons.

Table (2) also gives us a snapshot into the disparities in inflation amongst countries. For example, Portugal, Sweden, and Ireland appear to have the most volatility in inflation compared to the other countries, and quite notably, France, Austria, Denmark, and the Netherlands appear to exhibit the least variation in realized inflation. From the RMSE calculations, forecasters from Switzerland and Canada appear to display the highest average predictive accuracy at almost every forecast horizon.

In the section which follows, I use the data-sets discussed above to conduct four well-known tests of full-information rational expectations and evaluate the associated findings.

3 Evidence of Heterogeneous 'Mistakes'

In this section, I conduct four well-known tests for full-information rational expectations on the cross-country consensus forecasts of inflation at varying forecast horizons. The objective is to highlight the similarities as well as the variation in the magnitude and direction of the results. Unsurprisingly, all countries display departures from FIRE, however, the results differ substantially. I begin with the Mincer-Zarnowitz (1969) test, followed by a simple Bias test. Both tests evaluate the forecaster's ability to make accurate predictions. I continue with the two efficiency tests – the Coibion-Gorodnichenko (2015) test, and the basic test for auto-correlation of forecast errors.

3.1 Mincer-Zarnowitz (1969) Test

This test is based on the premise that a forecast is unbiased if the realized value of a variable moves one-for-one with its forecast. In other words, forecasters should make perfect predictions. The test is specified as,

$$\pi^i_{t+h} = \alpha^i + \beta^i F_t \pi^i_{t+h} + u^i_{t+h}, \tag{1}$$

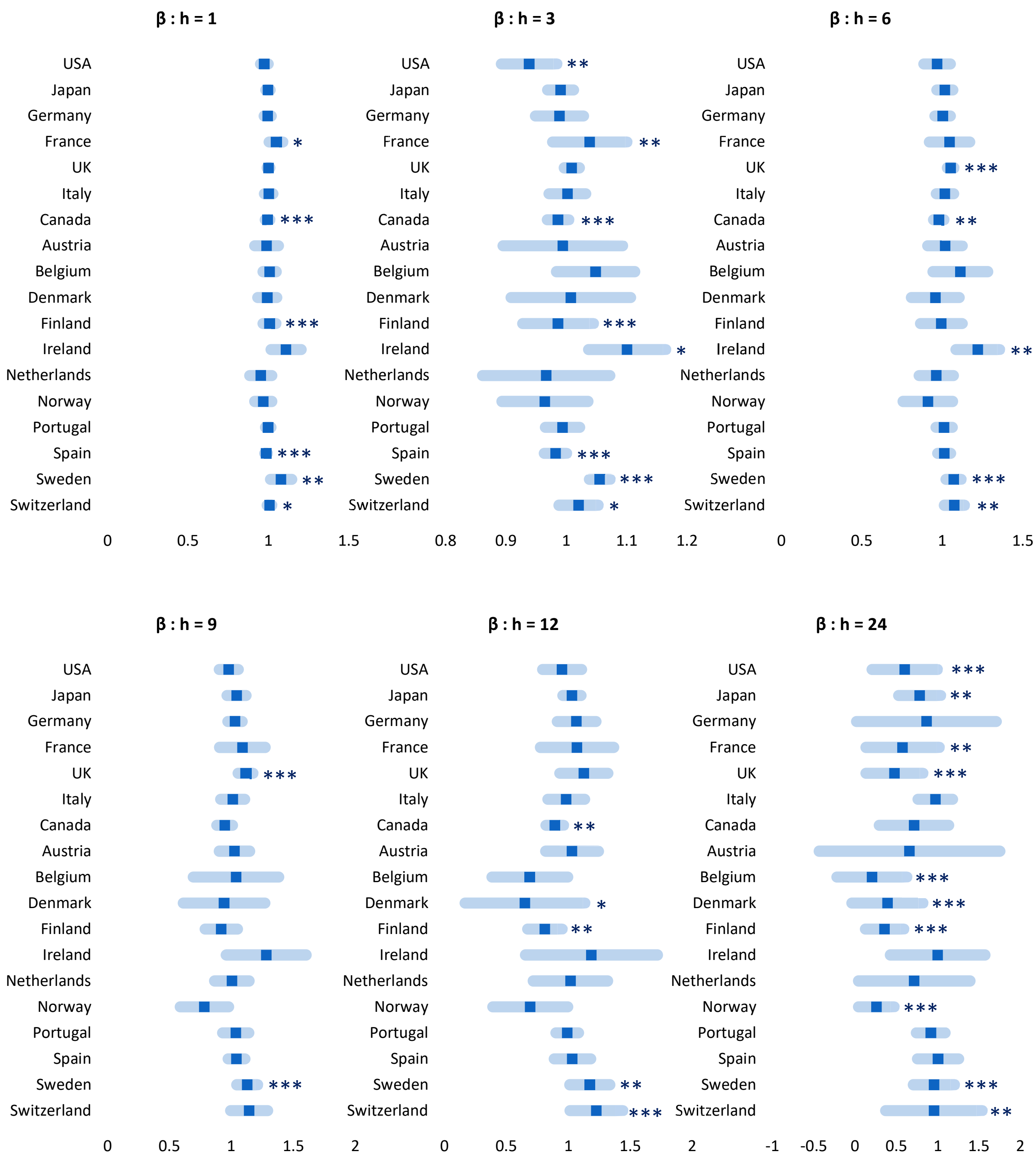
where, the left-hand-side variable is the realized inflation of country i at time $t{+}h$, α^i is the country-specific intercept, $F_t\pi^i_{t+h}$ is the h-period-ahead inflation forecast made at time t, and u^i_{t+h} is the rational expectations error term. Under full-information rational expectations, the null hypothesis is $\alpha^i = 0$, that is, forecasts should exhibit no bias and $\beta^i = 1$. Figure (2) shows the estimated beta coefficients for all countries from regression specification (1) at selected forecast horizons.

These results confirm the rejection of FIRE at every forecast horizon, but more so at the most distant forecast horizon, $h = 24$. One can also readily observe that within the cross-section of results, there is a positive relationship between realized and forecasted inflation. On average, forecasters seem to accurately predict the direction of realized inflation. This is not surprising, however, since predictions are made monthly.

What is unique about these findings is the disparity in the beta coefficients across forecast horizons. A simple calculation reveals that the average beta coefficient at forecast horizons, $h = 12, 24$ are 0.98 and 0.69, respectively. The message here is that at the furthest forecast horizon, $h = 24$, forecasters appear to produce very conservative inflation forecasts since there is less than a one-for-one relationship between predicted and realized inflation. That is, they foresee inflation being significantly higher than realized. At horizon, $h = 12$, this result continues on average, but much less so, and as the forecast horizons decrease, forecasts become more in line with realized inflation but still exhibit bias.

Amongst the countries, instances of under- and over-prediction of inflation are also seen across the horizons. For example, at horizon $h = 3$, forecasters in the US continue to over-predict inflation, whereas forecasters in both Sweden and Switzerland under-predict this variable, all with statistical significance. Over the same horizon, forecasters in France and Ireland appear to produce much less conservative predictions than all countries. Finally, through all horizons, forecasters in Switzerland and Sweden appear to consistently under-predict inflation.

Figure 2. Mincer-Zarnowitz Test



Note: Figure (2) shows the estimated beta coefficient (dark blue points) from the Mincer-Zarnowitz Test at forecast horizons, 1, 3, 6, 9, 12, and 24 months. Shaded bars represent the 95% confidence intervals. For each country n = 31 *p < 0.1, **p< 0.05, ***p< 0.01.p-values computed using Newey-West standard errors with automatic lag selection.

What becomes clear from the results of this test is that forecasters' view of realized inflation differs across forecast horizons, but more so across countries.

3.2 Bias Test

Mankiw and Shapiro (1986) [22] have challenged the classic Mincer-Zarnowitz (1969) test on the premise that the auto-regressive nature of the right-hand-side variable leads to a small sample bias which causes the null hypothesis to be rejected too often, in favor of the conclusion that the forecast is biased. Relatedly, Cukierman et. *al.* (2020) [16] show that the presence of serial correlation in the error term also causes the null hypothesis to be rejected too often in favor of the conclusion of biased forecasts.

To address these issues, Equation (1) is modified, such that: $e^i_{t+h|t} = \pi^i_{t+h} - F_t\pi^i_{t+h}$, where, $e^i_{t+h|t}$, is the one period ahead forecast error. The formal Bias test can then be specified as the following zero mean test:

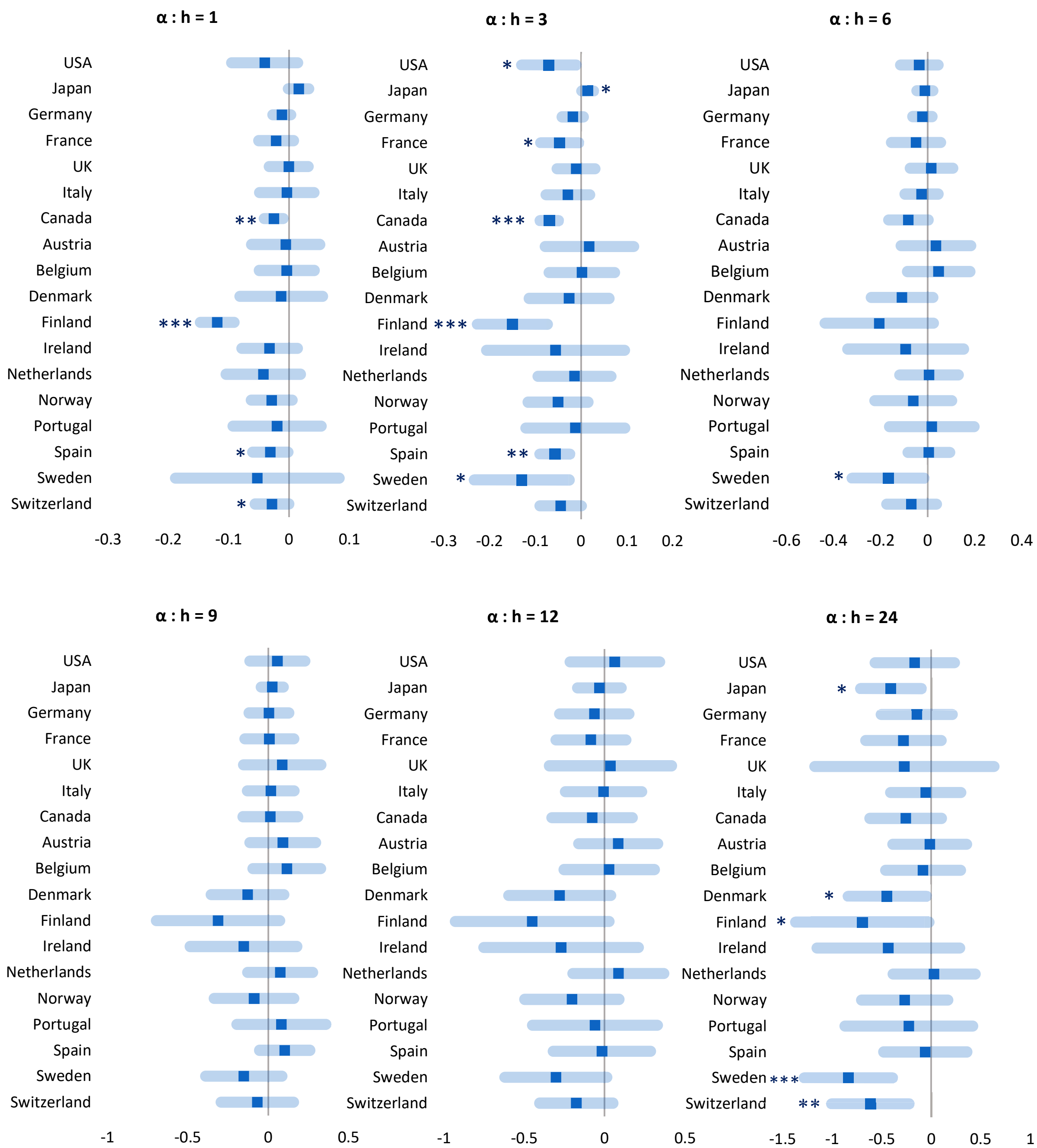
$$e^i_{t+h|t} = \alpha^i + u^i_{t+h}. \tag{2}$$

Similar to the Mincer-Zarnowitz test, the Bias test is based upon the premise that under full-information rational expectations, agents should on average forecast the variable of interest precisely, with zero prediction error. That is, forecasts should be unbiased. Therefore, the null hypothesis of this test is $\alpha^i = 0$. An $\alpha^i > 0$ indicates a downward bias or under-prediction of inflation, whereas an $\alpha^i < 0$ indicates an upward bias or over-prediction of inflation. Figure (3) shows the estimated α^i coefficients from specification (2) at selected forecast horizons for each country.

While these results indicate that forecasts are unbiased at almost every horizon, there are some useful observations that can be drawn. Firstly, forecasters appear to over-predict inflation across almost every horizon. Consistent with the results from Figure (2) The findings from equation (1), and consistent with the literature on departures from full-information rational expectations, the results indicate the presence of bias in the predictions of professional forecasters across countries. The result also reveals slightly different, yet consistent findings when compared to those from specification (2). For example, at the furthest forecast horizon, $h = 24$, forecaster 'mistakes' appear to be less prevalent. In fact, the results show that at the $h = 12$ forecast horizon, forecasts exhibit no statistically significant bias.

At the shortest forecast horizon, however, similar to the results in Figure (2), while forecasts become more in line with full-information rational expectations, there is significant heterogeneity in the results. Specifically, forecasts from the USA, Japan, France, Canada, Spain, Sweden and Finland appear to be statistically different from zero. Further, with the exception of Japan at horizon $h = 3$, all statistically significant 'mistakes' appear to be in the direction of over-prediction of inflation or upward-biased estimates, that is, $\alpha^i < 0$.

Figure 3. Bias Test



Note: Figure (3) shows the estimated intercept (dark blue points) from the Bias Test at forecast horizons, 1, 3, 6, 9, 12, and 24 months. Shaded bars represent the 95% confidence intervals. For each country n = 31 *p < 0.1, **p< 0.05, ***p< 0.01.p-values computed using Newey-West standard errors with automatic lag selection.

3.3 Coibion-Gorodnichenko (2015) Test

Under full-information rational expectations, since forecasters instantaneously incorporate all new and relevant information into their models, a key implication is that forecast errors should be unpredictable from any information available at time t or earlier. This is the basic premise of the Coibion-Gorodnichenko (2015) test, which is now widely used in the literature. The general specification of the test is:

$$\pi^i_{t+h} - F_t\pi^i_{t+h} = \alpha^i + \beta^i(F_t\pi^i_{t+h} - F_{t-1}\pi^i_{t+h}) + u^i_t, \tag{3}$$

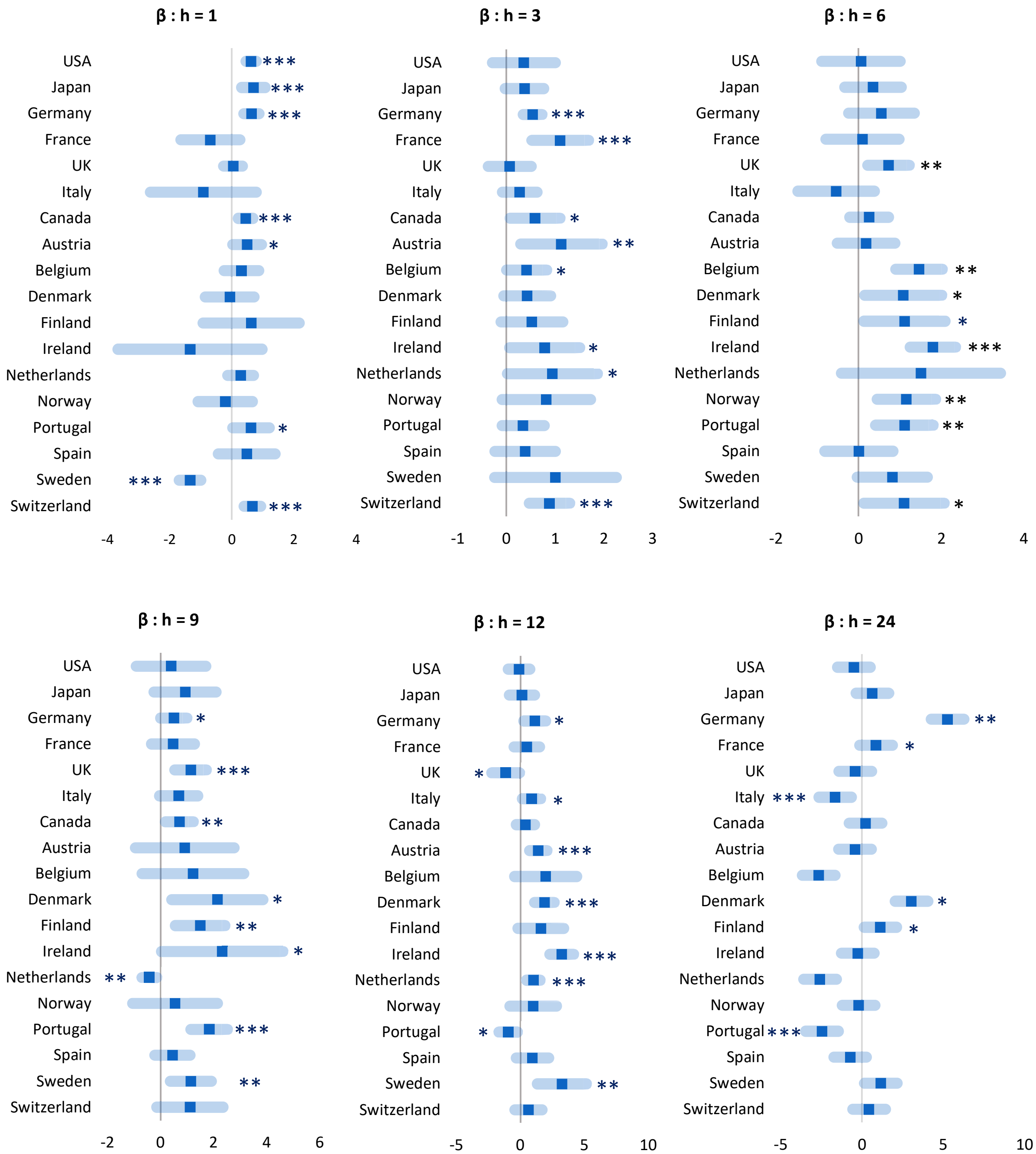
where, the left hand side of equation (3), $\pi^i_{t+h} - F_t\pi^i_{t+h}$, represents the one period ahead forecast error. This term is regressed against the time t forecast revision, $(F_t\pi^i_{t+h} - F_{t-1}\pi^i_{t+h})$, for each country i. This forecast revision measures the forecasters' reaction to relevant news on the economy which they have processed. The final term u^i_t is the rational expectations error which is uncorrelated with all information received at time t or earlier.

Equation (3) tests the null hypothesis of $\alpha^i = 0$ and $\beta^i = 0$. That is, forecast errors should be both unbiased and unpredictable. Figure (3) displays the estimated beta coefficient by forecast horizon. The novelty of this test is that the value of β^i also maps the extent of under- or over-reaction to new information the agent receives. This test therefore provides an assessment on the role of information frictions in deviations from full-information rational expectations. Coibion and Gorodnichenko (2015) [13] argue that a $\beta^i > 0$ implies that on average agents are slow to update their information sets given a macroeconomic shock or new information on the state of the economy, which indicates under-reaction. Conversely, $\beta^i < 0$ implies agents' over-reaction to new news. Therefore, this test provides an assessment on the role of information frictions in deviations from full-information rational expectations. Within the empirical literature, there is strong evidence to support the view that under-reaction is the predominant feature of consensus forecasts, while over-reaction characterizes individual forecasts.

The results in Figure (4) confirm cross-country departures from full-information rational expectations for most countries in the data-set with under-reaction as the predominant feature of these expectations. A simple calculation shows that the average beta coefficient at forecast horizons $h = 3, 6, 12, 24$ is, 0.59, 0.73, 0.87, and 0.03, respectively. Coibion and Gorodnichenko (2015) [13] show that based on the sticky-information model put forward by Mankiw and Reis (2002)[30], these coefficients can be mapped to the average duration of forecasters' information updates, $\hat{\lambda}$, where, $\hat{\lambda} \approx \dfrac{\hat{\beta}}{1+\hat{\beta}}$. Applying this formula, implies that on average forecasters across these 18 countries update their information sets, every 4.5, 5, 5.6 and .35 months, at forecast horizons, $h = 3, 6, 12, 24$, respectively.

These results also demonstrate a significant amount of heterogeneity across-countries. For example, not only is the range in the magnitude of the variation in the beta coefficients larger at the furthest forecast horizon, but interestingly, in countries such as Sweden, Italy, UK, Netherlands and Portugal I find evidence of over-reaction in aggregate expectations at forecast horizons $h = 1, 9, 12, 24$, respectively. This is in direct contradiction with the literature on aggregate expectations, particularly for inflation.

Figure 4. Coibion-Gorodnichenko Test



Note: The table shows the estimated beta coefficient (dark blue points) from the Coibion-Gorodnichenko Test at forecast horizons, 1, 3, 6, 9, 12, and 24 months. Shaded bars represent the 95% confidence intervals. For each country n = 30 *p < 0.1, **p< 0.05, ***p< 0.01.p-values computed using Newey-West standard errors with automatic lag selection.

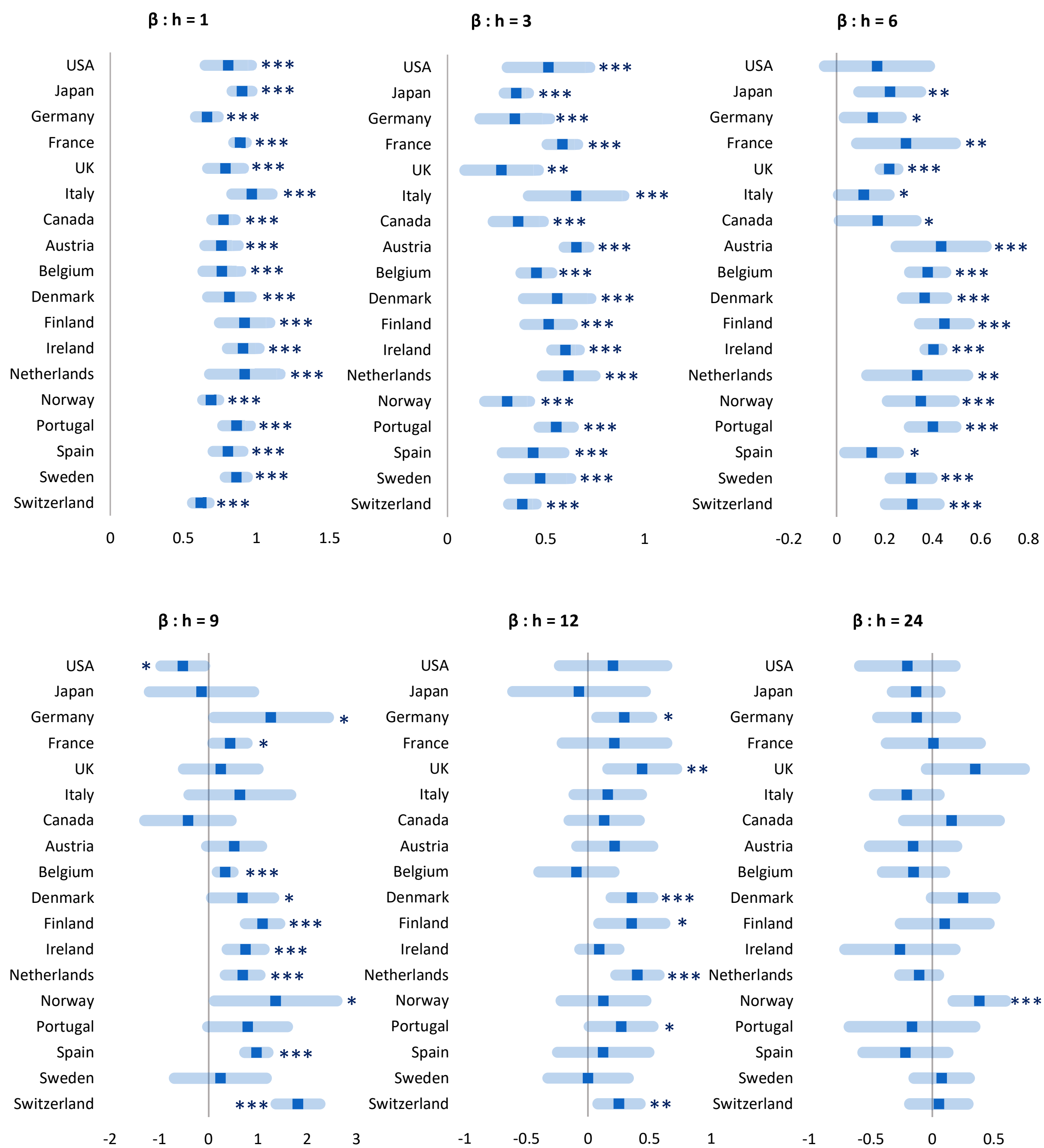
3.4 Auto-correlation of Forecast Errors

For a final test of forecaster 'mistakes', the following regression specification for auto-correlated forecast errors was conducted:

$$e^i_{t+h|t} = \alpha^i + \beta^i e^i_{t|t-h} + u^i_{t+h}. \tag{4}$$

In this specification, the h-period ahead forecast error, $e^i_{t+h|t}$, is regressed against its own past value h periods earlier, $e^i_{t|t-h}$. Similar to the Coibion-Gorodnichenko (2015) test, this specification is based upon the premise that the forecast contain all information available at time t or earlier, therefore, forecast errors should be serially uncorrelated. The null hypothesis is therefore, $\alpha^i = 0$ and $\beta^i = 0$. The results of specification (4) are shown in Figure (5) below.

Figure 5. Auto-correlation Test



Note: The table shows the estimated beta coefficient (dark blue points) from the Auto-correlation Test at forecast horizons, 1, 3, 6, 9, 12, and 24 months. Shaded bars represent the 95% confidence intervals. For each country n = 31 at h = 1, h = 3 and h = 6, n = 30 at h = 9 and h = 12, and n = 29 at h = 24. *p < 0.1, **p< 0.05, ***p< 0.01.p-values computed using Newey-West standard errors with automatic lag selection.

A clear pattern emerges from these results. Firstly, as with the other test results presented in this section, cross-country forecasts in this data-set appear to systematically depart from full-information rational expectations. This test in particular reveals significant positive serial correlation in forecaster predictions. Given the nature of the forecasts, however, it is not surprising that forecasts are serially correlated at the shortest forecast horizons , $h = 1$, 3, 6 and less so at the further horizons, $h = 12$, 24.

3.5 Summary of Findings

The results of the four tests presented in this section provide consistent empirical micro-evidence that cross-country inflation expectations of professional forecasters systematically depart from the assumptions of full-information rational expectations. While this over-arching conclusion has been the subject of many empirical papers which apply survey data, what is unique about this paper is the opportunity to observe cross-sectional heterogeneity in the forecaster 'mistakes'. What follows is a discussion on how these results may be interpreted and the associated implications.

As discussed in Section (2) approaches to analyzing forecaster 'mistakes' are often categorized into two broad areas within the literature. One approach is to ascribe these departures to the presence of constraints on the information processing capacity of forecasters. In such instances, models of sticky or noisy information are just some of the models used to explain the observed results. The second approach is to explain these departures with models of slow learning, the assumption is that the agent does not observe the true data generating process (DGP) at the time the new data is received, instead it is observed over time.

An interesting finding common to all tests, is that for some countries forecaster 'mistakes' or violations of full-information rational expectations occur at the nearest forecast horizon, $h = 3$. This result is at odds with this literature in two distinct ways. Firstly, the fact that these 'mistakes' occur at the shortest horizon when the forecaster has had the opportunity to observe up to nine months of realized inflation is suggestive of the fact that neither the presence of information constraints nor the assertion of incomplete knowledge of the DGP may be adequate explanations.

Secondly, these predictions are made by professional forecasters who are arguably less information-constrained than other agents. From Figure (1) and Table (1) we have observed that as early as forecast horizon $h = 6$, forecasters appear to make predictions that are very close to the realized values. Further, the results of the auto-correlation test from regression specification (4) show that at earlier forecast horizons, forecasts are highly serially correlated. Therefore, it is likely that the forecaster 'mistakes' observed for some countries may reflect the use of private information, judgement or other more country-specific factors considered by the forecaster which are not captured by these tests.

To confirm this assertion I turn to the results of the four tests and show that they all convey a similar message. The results of the Mincer-Zarnowitz test indicate a positive relationship between cross-country inflation expectations and the realized value of inflation at all horizons. The Bias test, which is a slight modification of the Mincer-Zarnowitz test, presents results which follow a similar theme. For all forecast horizons, we observe that forecasters systematically over-predict inflation. Together, the results of both tests imply that over the period of this study, the positive relationship between inflation and inflation expectations is in the direction of over-estimation or a more pessimistic view of future inflation compared to the realized values. This finding is in fact confirmed for the US in a recent study using CPI data from Survey of Professional Forecasters (SPF) and University of Michigan's Survey of Consumers (SOC) (see Bianchi et *al.* (2022) [6]).

Continuing to the results of the Coibion-Gorodnichenko (2015) test, one can observe that cross-country inflation expectations are subject to information frictions and overwhelmingly exhibit under-reaction. Given the results of the three aforementioned tests, an intuitive question is: Can forecasting inertia coexist with over-prediction in inflation? A likely line of reasoning is that forecasters, in aggregate, rely on their own private information or judgement or some other unexplained factors, to the extent that their views of inflation are little altered with new information. This is somewhat confirmed with the results of the auto-correlation test emphasizing the highly positive serial correlation in cross-country forecast errors, confirming forecasters' reliance on recent values as opposed to new, incoming information.

I conclude this section by highlighting the significant cross-sectional heterogeneity in the results of all four tests. No where is this more clearly seen than in the results of the Coibion-Gorodnichenko (2015) test. Despite the general consensus within the empirical literature that aggregate expectations exhibit under-reaction, the results of this test show clear instances of statistically significant over-reaction to new information by professional forecasters at the twelve and twenty-four month horizons for Italy, Portugal and the UK.

What makes this finding particularly interesting is that at these forecast horizons, forecasters have no prior data-points on realized inflation for the years being forecasted over, compared to the shorter forecast horizons. It is well-known from the literature (see Bordalo et *al.* (2020) [7] and Afrouzi et *al.* (2020) [1]) that over-reaction is a feature of individual forecasts and varies with the persistence of the variable being predicted. In these papers, the authors conclude that the less persistent the series, the stronger the over-reaction.

The key insight here is that departures from full-information rational expectations across-countries, while they share many common features, are heterogeneous in many respects. Therefore, the tendency to ascribe broad conclusions about the characteristics of expectations at the aggregate level, such as over- versus under-reaction or upward versus downward bias, without a comprehensive cross-sectional approach, is likely premature and misleading. The more serious implication is that our understanding of the expectation formation process is incomplete and the associated models and policy prescriptions inadequately considered.

In the section which follows I discuss the coexistence of common and heterogeneous factors which may impact the expectation formation process by focusing on the cross-country forecast errors.[8]

4 A Bayesian Dynamic Factor Model

Given the evidence of departures from full-information rational expectations across countries, I now examine the extent to which forecaster 'mistakes' may be co-move across countries. To the extent that co-movement exists and persists is suggestive that these departures from full-information rational expectations may play a role in cross-country inflation expectations.

To conduct this analysis I turn to a widely used class of models which decomposes a macroeconomic time series variable into components which are not directly observable by the econometrician - that is, *unobserved components models*. Included within this class of models are *state space models* where the observed time series depends linearly on a possibly unobserved state driven by a stochastic process. *Factor models* which distill information in large data-sets to a lower-dimensional set of (unobserved) factors, also fall within the class of state space models and can be modeled in a linear state space representation. In these models it is assumed that a set of n observed variables depends linearly on m unobserved common factors and on an individual or idiosyncratic component.[9]

This paper presents an empirical application of a *dynamic* factor model. A distinguishing feature of this model compared to a *static* factor model is that all inter-temporal cross-correlations among variables is accounted for by the dynamic unobserved factors.

4.1 Econometric Framework

Let $e_t = [e_{1t}, e_{2t}, ..., e_{nt}]', t = 1, ..., T, i = 1, ..., n$ denote a stationary $n \times 1$ vector of observable cross-country inflation forecast errors standardized to mean zero and unit variance at time t. Here, e_t is the difference between realized inflation at time t and predicted inflation at time *t-1* or $\pi_t - \pi^e_{t-1}$. A general representation of the factor model is given by:

$$e_t = \lambda f_t + u_t, \tag{5}$$

where, f_t is an $m \times 1$ vector of unobserved common factors which are believed to capture the co-movement of the cross-country panel of forecast errors and λ is an $n \times m$ matrix of unknown factor loadings which determine the exact linear combination of the m latent factors. Here, $\chi_t = \lambda f_t$ is referred to as the common component and $u_t = [u_{1t}, u_{2t}, ..., u_{n,t}]'$ as the idiosyncratic component, or country-specific noise of the model. It is assumed that $u_t \sim iid\ N(0, \Sigma)$ where Σ is set to be a diagonal matrix. This implies that $Eu_{it}u_{jt-s} = 0$ for $i \neq j$ or that the idiosyncratic

[8] Within the literature, there have been several papers discussing the endogenous link between inflation expectations and forecast errors. For example, Carvalho et *al.* (2022) [9] present a model of anchored expectations driven by the endogenous link between inflation expectations and forecast errors.

[9] Where $m < n$.

shocks are uncorrelated. Therefore, u_{jt} is a shock idiosyncratic to equation i and e_{it} is said to follow an *exact* factor model since all co-movements in the data arise from the latent factors plus an idiosyncratic shock.

Note that the above model is static since it does not allow for autocorrelation in the factors. For time series applications, however, the factors may be serially correlated and follow an autoregressive process of order q,

$$f_t = \Phi_1^f f_{t-1} + ... + \Phi_q^f f_{t-q} + \eta_t^f, \tag{6}$$

where, Φ_i^f is the $m \times m$ coefficient matrix corresponding to the ith lag of factor f_t and $\eta_t^f = [\eta_{1t}^f, \eta_{2t}^f, ..., \eta_{n,t}^f]'$ is the error term. Note here that $\eta_t^f \sim iid\ N(0, \Sigma^f)$, where Σ^f is assumed to be diagonal, that is $E\eta_{it}^f \eta_{it-s}^f = \sigma_f^2$ for $s = 0; 0$ otherwise.

Similarly, the evolution of the idiosyncratic errors may also follow an autoregressive process of order p,

$$u_t = \Phi_1 u_{t-1} + ... + \Phi_p u_{t-p} + \eta_t, \tag{7}$$

where, Φ_i is a matrix of autoregressive coefficients corresponding to the ith lag of u_t and the innovations, similar to the factor innovations, are assumed to be zero mean, contemporaneously uncorrelated normal random variables, $\eta_t \sim iid\ N(0, \Sigma)$, where Σ is set to be a diagonal matrix. Finally, note here that $E\eta_{it}^f \eta_{it-s} = 0\ \forall\ i, s$. That is, the idiosyncratic terms of the latent factor and the observables are contemporaneously uncorrelated normal random variables.

This allowance for serial correlation in the idiosyncratic errors and the latent factors is what makes the model dynamic. Together, equations (5) through (7) represent the dynamic factor model used to evaluate the inter-temporal co-movement of forecast errors across countries.[10] In the context of the linear Gaussian state space model, equation (5) represents the measurement equation and equations (6) and (7) the transition equations.

4.1.1 Identification Restrictions

It is well known that the dynamic factor model presented in equations (5) through (7) is not identifiable without further restrictions. It can be shown that by pre-multiplying the dynamic factor by any arbitrary full-rank $m \times m$ matrix defines a new model which is observationally equivalent to the original model. For example, for any non-singular matrix C, $\lambda f_t = \lambda C C^{-1} f_t$, therefore, a model with factors f_t and factor loadings λ will give the same fit as one with factors $C^{-1} f_t$ and factor loadings λC. As a result of this normalization issue, neither the signs nor the *scale* of the dynamic factors and the associated factor loadings are separately identified.

Two additional restrictions are required for unique identification. Firstly, it is common to set the variance of η_t^f to the identity matrix, as in Bai and Wang (2015) [5]. That is,

[10] This is the standard model put forward by Stock and Watson (1989) [46] and applied in the context of international business cycles by Kose et *al.* (2003) [27].

$$var \begin{bmatrix} \eta_{1t}^{f} \\ \eta_{2t}^{f} \\ \vdots \\ \eta_{nt}^{f} \end{bmatrix} = I_m.$$

This restriction addresses the *scale* issue. The second normalization restriction addresses the sign issue by setting λ to be a lower triangular matrix with ones on the diagonal.[11] Appendix A details the mathematical proof that these two restrictions are sufficient to uniquely identify the dynamic factor model presented in equations (5) through (7).

4.1.2 Selecting the Number of Dynamic Factors

Prior to estimation of the dynamic factor model, an important consideration is the number of common latent factors to be included. This decision is largely approached via optimization of an *Information Criterion* to consistently estimate the number of factors. The objective is often to select the smallest number of latent trends without loss of too much information.

Within the empirical literature, among the most widely used approaches to factor selection are the Bai and Ng (2002) [3] criterion for static factor models where the minimization of the information criteria is based on a trade-off between the quality of the adjustment of the model to the data and the risk of over-adjustment.

In the context of dynamic factor models, the Bai and Ng (2007) criteria [4] builds on the Bai and Ng (2002) [3] criteria by first taking the optimal number of factors selected as given. The authors then estimate a VAR*(p)* on these factors using the Bayes Information Criterion (BIC) and then as a final step, use the Bai and Ng (2007) [4] criteria to obtain the optimal number of dynamic factors.

Finally, Stock and Watson (2005) [45] and Amengual and Watson (2007) [2] show that the Bai and Ng (2002) [3] criteria can in fact be used to estimate the number of dynamic factors.

4.1.3 Estimation

Following Stock and Watson (1989, 1993) [44, 46] the classical approach to this system of equations has been to apply statistical techniques which make use of the Kalman filter to estimate model parameters and the Kalman smoother to extract an estimate of the latent factor, however, there are a few drawbacks to this method. Firstly, when n is large, this method can be computationally demanding and secondly, since the factors are random variables, when n is finite, estimators will not be consistent.

[11] There are several variations to addressing the normalization issue in this model. For example, to address the *scale* issue, Sargent and Sims (1977) [40] and Stock and Watson (1989, 1993) [44, 46] assume that σ_f^2 is equal to a constant, which is quite similar to the approach taken in this paper. Further, it is common to identify the signs by requiring one of the factor loadings to be positive as in Otrok and Whiteman (1998) [38] or to require λ to be a lower triangular matrix with strictly positive diagonal elements as specified in Bai and Wang (2015) [5].

In the more recent literature (see Otrok and Whiteman (1998) [38]), these equations have also been estimated using Bayesian inference which makes use of *posterior simulation*. Here, a Monte Carlo Markov chain (MCMC) procedure can be employed to generate samples from a given target distribution. In the empirical application of this paper, the distribution of interest is the *joint posterior distribution* of the model parameters and latent factors given the observed data, that is, $p(\Theta, f|\text{data})$. Here Θ represents the parameters, $\lambda, \Sigma, \Sigma^f, \phi_i^f, i = 1, ..., q$, and $\phi_i, i = 1, ..., p$.

Given the large cross-sectional data-set used in this paper, however, sampling from the joint posterior distribution directly is complex and even prohibitive. The Gibbs sampler, a MCMC sampling algorithm, addresses this problem by first expressing this distribution in terms of it's complete set of conditional distributions and then iteratively sampling from the individual *posterior conditional distributions*.[12] Casella and George (1992) [10] show that under weak conditions, this iterative sampling converges to drawing from the joint posterior distribution directly as the number of sampling draws, $j \to \infty$. With this approach the complexity of sampling directly from the joint posterior distribution is simplified via a series of smaller more manageable problems.[13]

As an example of how the Gibbs sampling methodology cycles through conditional distributions, starting with an initial (specified) value, f^0, a sample is drawn from the posterior distribution of the parameters conditional on the factor, $p(\Theta|f^0, data)$, which produces a drawing, Θ^1. Next a sample is drawn from the distribution of the latent factor, conditional on the parameters, $p(f|\Theta^1, data)$ which produces a drawing, f^1. This completes one Gibbs sample or the first step of the *Markov Chain*.[14]

Continuing this process generates a *Gibbs sequence* of random variables, $f^0, \Theta^1, f^1, ...\Theta^j, f^j$. That is, once the initial value f^0 is specified, the remainder of the sequence is obtained iteratively by alternately generating, at each step, drawings from,

$$\begin{aligned} \Theta^j &\sim p(\Theta|f^{j-1}, data) \\ f^j &\sim p(f|\Theta^j, data), \end{aligned} \tag{8}$$

for $j = 1, ..., J$. Once the samples are obtained, the sample mean and median of the posterior distributions represent the traditional point estimates, while the percentiles can be used as confidence intervals.

[12] Based on the assumption that the joint distribution can be characterized by its complete set of conditional distributions. Usually applied to standard continuous distributions such as Normal, t, Beta or Gamma or discrete distributions such as Binomial, Multinomial or Dirichlet.

[13] Chan et *al.* (2019) [11] present a comprehensive introductory example of the Gibbs sampling routine.

[14] Note that the order of sampling, that is, which conditional is sampled first or second does not matter, what matters is that a sample is drawn from each of the unknown blocks. In the application presented in this paper, there are two blocks (parameters and the cross-country factor). Kose et al. (2003) [27] present an application of the Gibbs sampler with four blocks in their application of dynamic factor analysis to international business cycles with a multi-level factor structure.

5 The Cross-Country Dynamic Factors in Forecaster 'Mistakes'

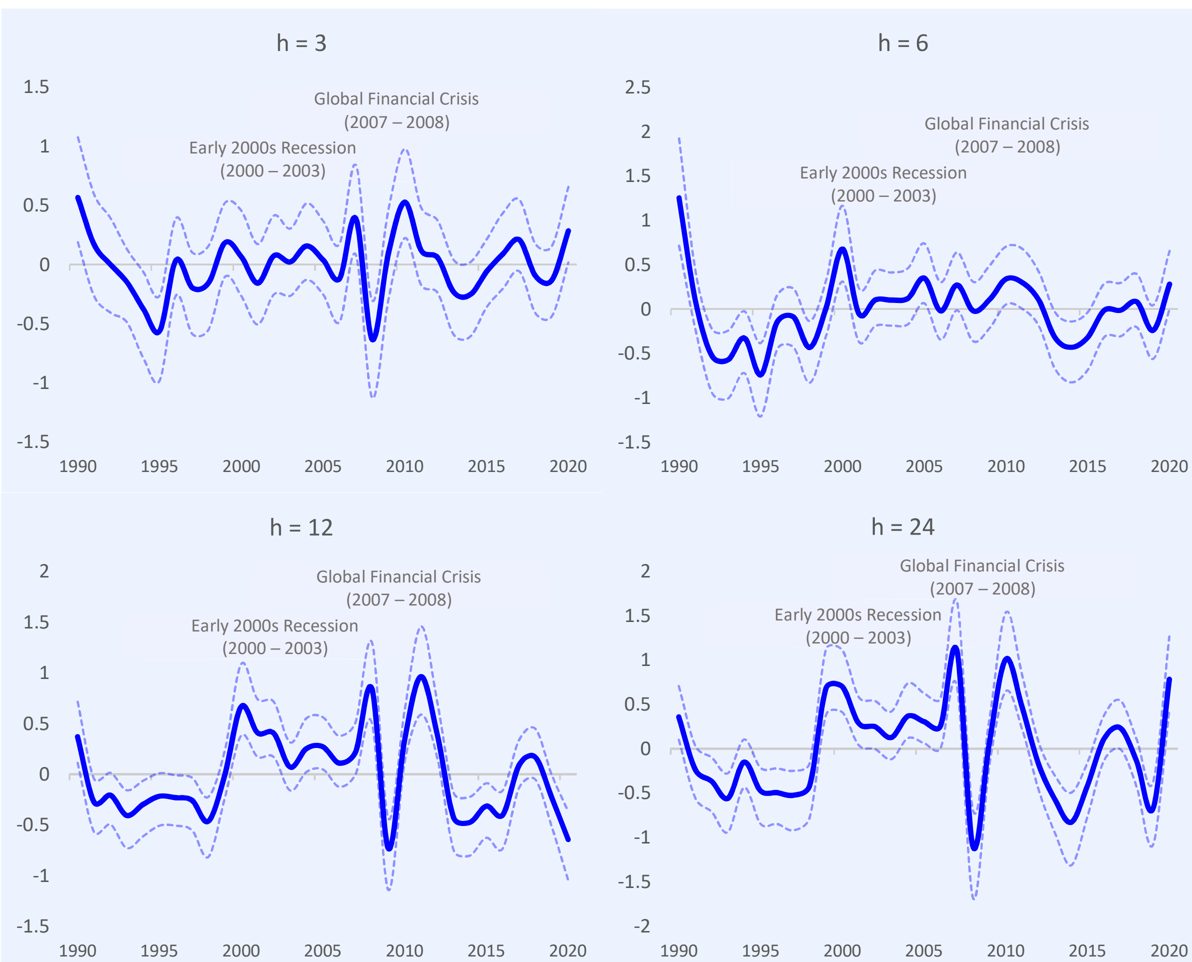
Figure (6) presents the median posterior distribution of the first latent factor in cross-country expectational errors estimated from equations (5) through (7) that is,

$$e_t^h = \lambda^h f_t^h + u_t^h, \tag{9}$$

where $h = 3, 6, 12, 24$ month forecast horizons, and f_t^h and u_t^h follow autoregressive processes of orders $q = 1$ and $p = 1$, respectively. The fluctuations in the factor indeed reflect the major economic events over the period of this study, such as the recession of the early 2000s which affected advanced economies and the 2007 - 2008 Global Financial Crisis (GFC). The factors, however, also appear to only be statistically significant at the $h = 24$ and $h = 12$ forecast horizons.

Figure 6. Cross-Country Dynamic Factor in Forecast Errors

Horizon Analysis



Note: The chart shows a time series of dynamic latent factors of inflation forecast errors for 18 OECD countries at forecast horizons of 3, 6, 12 and 24 months. Light blue broken lines represent the 95% confidence interval. Forecast errors have been standardized to have mean zero and unit variance.

In three of the four graphs ($h = 3, 12, 24$) the narrowing of the bands around these events is likely indicative of the increased relevance of the cross-country factor. This assertion is in line with Borio and Filardo (2007) [8] who suggest that the relevance of global inflation is time-varying. Moreover, this narrowing also seems to arise in the 2019 - 2020 period marked by the COVID-19 pandemic, particularly at the $h = 24$ forecast horizon. An interesting extension of this paper is to assess the increased relevance of this factor coming out of the pandemic, as more data become available.

Another observation from these graphs is that forecasters' predictions oscillated between under-prediction of inflation (positive latent factor) and over-prediction (negative latent factor) from the very beginning of the period of study, as developed economies recovered from 'The Great Inflation' which ended in 1982, to the period leading up to and emerging from the GFC, at most forecast horizons.

From the theoretical literature, it is well-known that expectational errors can be ascribed to reasons such as the presence of information frictions as in Mankiw and Reis (2002)[30], Woodford (2002) [47] and Sims (2003) [42], rational inattention as in Gabaix (2014) [20] and the existence of extrapolative tendencies (Gennaioli et *al.* (2015)[21]) which is closely related to the well-known representativeness bias in human judgment put forward by Kahneman and Tversky (1972) [25]. Within the more recent literature, and in the context of the professional forecaster, these errors are ascribed to the forecasters' incomplete knowledge of the data generating process which they learn slowly over time as in Farmer et *al.* (2021) [18].

While the presence of a global factor reflects commonalities among forecasters across-countries, another interesting issue is the degree of heterogeneity or the country-specific fundamentals which determine the interaction with the world factor. I begin with a measure of the relative contribution of the cross-country factor to variations in the forecast errors in each country. Essentially, I decompose the variance of each observable into the fraction that is due to the cross-country factor (or common unobserved component) and the fraction driven by the idiosyncratic component. In an approach similar to Kose et *al.* (2003), the variance of the observable forecast errors due to the cross-country factor i can be written as:

$$var(e_{it}^h) = (\lambda_i^h)^2 var(f_t^h) + var(u_{it}^h), \tag{10}$$

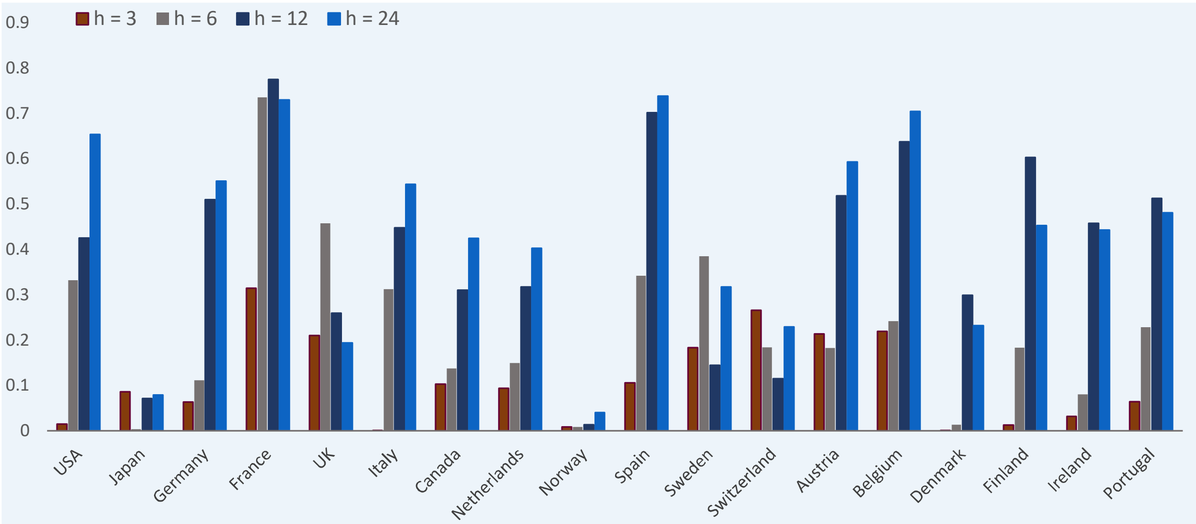
where the fraction of volatility attributed to the cross-country factor for each country is measured as,

$$var(f_{it}^h) = \frac{(\lambda_i^h)^2 var(f_t^h)}{var(e_{it}^h)} \ .$$

Figure (7) presents the variance shares attributable to the cross-country latent factors for each country for forecast horizons $h = 3, 6, 12, 24$.

Figure 7. Variance Decomposition of Cross-Country Forecast Errors

Horizon Analysis



Note: The chart shows the variance decomposition from equation (11) for 18 OECD countries at forecast horizons of 3, 6, 12 and 24 months. At each forecast horizon, the value of each bar represents the portion of the variation in the cross-country factor which accounts for the variation in the specific country factor.

Note here that the fraction of the variance of the observable (forecast errors) due to the idiosyncratic or country-specific component is simply $1 - var(e^h_{it})$. On average, the variations in the cross-country factor appear to explain a significant fraction of the fluctuations in each country's forecast errors at the furthest forecast horizons, $h = 24$, 12. At these horizons, the average contribution of the cross-country factor to the total variation amongst all countries is 0.43 and 0.40, respectively. At earlier forecast horizons, the role of the cross-country factor diminishes, with the average variance contribution of the factor declining to 0.23 and 0.11 at the $h = 6$ and $h = 3$ horizons, respectively. However, these generalizations do not hold for all countries. Notably, forecast errors from forecasters in France, Spain and Belgium, appear to be more correlated with the cross-country factor at $h = 24$, 12 than other countries.

This finding can be examined in conjunction with the results of the factor loadings presented in Table (3). We can think of these loadings as capturing developments specific to the domestic economies as opposed to the variance decomposition presented in Figure (7) which describes how the factor affects all countries. The Table shows that in general forecasters appear to load most heavily on the cross-country factor at the $h = 12$ forecast horizon and least at the $h = 3$ horizon. It is, however, interesting that forecasters from France, Spain and Belgium continue to load heavily on the cross-country as opposed to developments specific to their domestic economies.

A key insight from these results is that while the the evidence points to the existence of the cross-country latent factor in forecast errors there is significant heterogeneity across countries.It is likely that these differences are driven by factors, including behavioral factors, specific to forecasters in domestic economies. The next section of this paper brings together the findings of heterogeneity in departures from full-information rational expectations with the findings implied by the dynamic factor analysis presented.

Table 3. Factor Loadings

Horizon Analysis

	h= 24	h = 12	h = 6	h = 3
USA	1.87	1.72	1.58	0.38
Japan	0.82	0.80	0.17	0.87
Germany	1.74	1.85	0.99	0.76
France	1.95	2.24	2.23	1.51
UK	1.19	1.41	1.79	1.29
Italy	1.74	1.76	1.53	0.09
Canada	1.59	1.52	1.09	0.95
Netherlands	1.56	1.54	1.13	0.90
Norway	0.61	0.36	0.29	0.28
Spain	1.97	2.13	1.60	0.96
Sweden	1.43	1.10	1.67	1.21
Switzerland	1.26	1.00	1.23	1.40
Austria	1.79	1.87	1.21	1.30
Belgium	1.92	2.04	1.38	1.31
Denmark	1.27	1.49	0.35	0.09
Finland	1.63	1.99	1.23	0.34
Ireland	1.60	1.78	0.85	0.54
Portugal	1.66	1.86	1.34	0.76

Note: The table shows the factor loadings at forecast horizons of 3, 6, 12, and 24 months. From the dynamic factor model specified in equations (6) through (8), these loadings determine the exact linear combination of the latent cross-country factor for each country over the specified forecast horizon.

6 Conclusion

The findings outlined in Sections (3) through (5) present a consistent yet heterogeneous perspective on cross-country inflation expectations. The overwhelming message from the tests of rational expectations, is that while forecasters across countries make predictions which violate the assumptions of full-information rational expectations they do so in a distinctively heterogeneous manner, to the extent that traditional models which account for these departures are unable to sufficiently address the findings. In other words, there is no one-size-fits-all model when examining the expectation formation process across countries. The results of the dynamic factor analysis admit a similar story. We observe compelling evidence of the existence of a cross-country latent factor in forecast errors at both the $h = 12$ and $h = 24$ forecast horizons, however, upon close examination, there is significant heterogeneity driving these results.

Undoubtedly, over the years, there has been significant progress in our understanding of the expectation formation process and in particular how we account for departures from full-information rational expectations. As discussed throughout this paper, models of sticky information and inattentiveness (Mankiw and Reis, (2002, 2010)) [30] [31], rational inattention (Sims (2003)) [42], adaptive learning, (Evans and Honkapohja, (2001))[17], memory (Malmendier and Nagel (2016)) [29] and diagnostic expectations (Bordalo et *al.* (2020)) [7] to name a few, have emerged as some of the leading approaches to accounting for forecaster 'mistakes'.

It is difficult, however, to overlook the fact that researchers have yet to settle on a unified approach to explaining key findings. This paper sits at the core of this issue. While we observe commonalities, the breadth and depth of the heterogeneity in findings often stand in the way of the broad adoption of our models. It is not surprising, therefore, that foundational macroeconomics classes still espouse the underlying tenets of full-information rational expectations as the bedrock upon which macroeconomic modeling is built.

Progress toward a more unified approach, perhaps in the form of a more parsimonious model to account for departures from full-information rational expectations, begins with the recognition of the need for flexibility in analyzing expectations. It also calls for an even more comprehensive analysis of the cross-section of expectations among agents and countries. This is the key insight of this paper and the premise upon which much of the analysis rests.

References

[1] Hassan Afrouzi et al. *Overreaction and Working Memory*. en. Tech. rep. w27947. National Bureau of Economic Research, Oct. 2020 (cit. on p. 19).

[2] Dante Amengual and Mark W Watson. "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel". In: *Journal of Business & Economic Statistics* 25.1 (Jan. 2007), pp. 91–96. ISSN: 0735-0015 (cit. on p. 22).

[3] Jushan Bai and Serena Ng. "Determining the Number of Factors in Approximate Factor Models". In: *Econometrica* 70.1 (2002). Publisher: [Wiley, Econometric Society], pp. 191–221. ISSN: 0012-9682 (cit. on p. 22).

[4] Jushan Bai and Serena Ng. "Determining the Number of Primitive Shocks in Factor Models". In: *Journal of Business & Economic Statistics* 25.1 (2007). Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 52–60. ISSN: 0735-0015 (cit. on p. 22).

[5] Jushan Bai and Peng Wang. "Identification and Bayesian Estimation of Dynamic Factor Models". In: *Journal of Business & Economic Statistics* 33.2 (2015). Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 221–240. ISSN: 0735-0015 (cit. on pp. 3, 21, 22, 33).

[6] Francesco Bianchi, Sydney C. Ludvigson, and Sai Ma. "Belief Distortions and Macroeconomic Fluctuations". en. In: *American Economic Review* (). ISSN: 0002-8282 (cit. on p. 19).

[7] Pedro Bordalo et al. "Overreaction in Macroeconomic Expectations". en. In: *American Economic Review* 110.9 (Sept. 2020), pp. 2748–2782. ISSN: 0002-8282 (cit. on pp. 2, 5, 19, 28).

[8] Claudio E. V. Borio and Andrew J. Filardo. *Globalisation and Inflation: New Cross-Country Evidence on the Global Determinants of Domestic Inflation*. en. SSRN Scholarly Paper. Rochester, NY, May 2007 (cit. on pp. 6, 25).

[9] Carlos Carvalho et al. "Anchored Inflation Expectations". en. In: *American Economic Journal: Macroeconomics* (). ISSN: 1945-7707 (cit. on p. 20).

[10] George Casella and Edward I. George. "Explaining the Gibbs Sampler". In: *The American Statistician* 46.3 (1992). Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 167–174. ISSN: 0003-1305 (cit. on p. 23).

[11] Joshua Chan et al. *Bayesian Econometric Methods*. en. ISBN: 9781108525947 Publisher: Cambridge University Press. Aug. 2019 (cit. on p. 23).

[12] Matteo Ciccarelli and Benoît Mojon. "GLOBAL INFLATION". In: *The Review of Economics and Statistics* 92.3 (2010). Publisher: The MIT Press, pp. 524–535. ISSN: 0034-6535 (cit. on p. 6).

[13] Olivier Coibion and Yuriy Gorodnichenko. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". In: *The American Economic Review* 105.8 (2015). Publisher: American Economic Association, pp. 2644–2678. ISSN: 0002-8282 (cit. on pp. 2, 5, 8, 15).

[14] Dean Croushore and Tom Stark. "A real-time data set for macroeconomists". In: *Journal of Econometrics* 105.1 (2001). Publisher: Elsevier, pp. 111–130. ISSN: 0304-4076 (cit. on pp. 2, 6).

[15] Dean Croushore and Tom Stark. "A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?" In: *The Review of Economics and Statistics* 85.3 (2003). Publisher: The MIT Press, pp. 605–617. ISSN: 0034-6535. (Visited on 08/16/2022) (cit. on p. 6).

[16] Alex Cukierman, Thomas Lustenberger, and Allan Meltzer. "The Permanent-Transitory Confusion: Implications for Tests of Market Efficiency and for Expected Inflation During Turbulent and Tranquil Times". en. In: *Expectations: Theory and Applications from Historical Perspectives*. Ed. by Arie Arnon, Warren Young, and Karine van der Beek. Springer Studies in the History of Economic Thought. Cham: Springer International Publishing, 2020, pp. 215–238. ISBN: 978-3-030-41357-6 (cit. on p. 13).

[17] George W. Evans and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, 2001. ISBN: 978-0-691-04921-2 (cit. on p. 28).

[18] Leland Farmer, Emi Nakamura, and Jón Steinsson. *Learning About the Long Run*. en. Tech. rep. 29495. Publication Title: NBER Working Papers. National Bureau of Economic Research, Inc, Nov. 2021 (cit. on p. 25).

[19] Adriana Z Fernandez, Evan F Koenig, and Alex Nikolsko-Rzhevskyy. "Real-Time Historical Dataset Enhances Accuracy of Economic Analyses". en. In: (2012), p. 4 (cit. on p. 6).

[20] Xavier Gabaix. "A Sparsity-Based Model of Bounded Rationality *". en. In: *The Quarterly Journal of Economics* 129.4 (Nov. 2014), pp. 1661–1710. ISSN: 0033-5533, 1531-4650 (cit. on p. 25).

[21] Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. *Expectations and Investment*. en. Tech. rep. w21260. Cambridge, MA: National Bureau of Economic Research, June 2015, w21260 (cit. on p. 25).

[22] N. Gregory Mankiw and Matthew D. Shapiro. "Do we reject too often?: Small sample properties of tests of rational expectations models". en. In: *Economics Letters* 20.2 (Jan. 1986), pp. 139–145. ISSN: 0165-1765 (cit. on p. 13).

[23] Lars Jonung and David E. Laidler. "Are Perceptions of Inflation Rational? Some Evidence for Sweden". en. In: *American Economic Review* 78.5 (1988). Publisher: American Economic Association, pp. 1080–1087. URL: https://ideas.repec.org/a/aea/aecrev/v78y1988i5p1080-87.html (visited on 11/18/2021) (cit. on p. 4).

[24] Robert E. Lucas Jr and Thomas J. Sargent, eds. *Rational Expectations and Econometric Practice, Volume 2*. English. 1st edition. Minneapolis: Univ Of Minnesota Press, Nov. 1981. ISBN: 978-0-8166-1071-6 (cit. on p. 4).

[25] Daniel Kahneman and Amos Tversky. "Subjective probability: A judgment of representativeness". en. In: *Cognitive Psychology* 3.3 (July 1972), pp. 430–454. ISSN: 0010-0285 (cit. on p. 25).

[26] Evan F. Koenig. "Is the markup a useful real-time predictor of inflation?" en. In: *Economics Letters* 80.2 (Aug. 2003), pp. 261–267. ISSN: 0165-1765 (cit. on p. 6).

[27] M. Ayhan Kose, Christopher Otrok, and Charles H. Whiteman. "International Business Cycles: World, Region, and Country-Specific Factors". en. In: *American Economic Review* 93.4 (Sept. 2003), pp. 1216–1239. ISSN: 0002-8282 (cit. on pp. 6, 21, 23).

[28] Robert E Lucas. "Expectations and the neutrality of money". en. In: *Journal of Economic Theory* 4.2 (Apr. 1972), pp. 103–124. ISSN: 0022-0531 (cit. on p. 4).

[29] Ulrike Malmendier and Stefan Nagel. "Learning from Inflation Experiences *". In: *The Quarterly Journal of Economics* 131.1 (Feb. 2016), pp. 53–87. ISSN: 0033-5533 (cit. on p. 28).

[30] N. G. Mankiw and R. Reis. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve". en. In: *The Quarterly Journal of Economics* 117.4 (Nov. 2002), pp. 1295–1328. ISSN: 0033-5533, 1531-4650 (cit. on pp. 15, 25, 28).

[31] N. Gregory Mankiw and Ricardo Reis. "Chapter 5 - Imperfect Information and Aggregate SupplyWe are grateful to students at Columbia University and Faculdade de Economia do Porto for sitting through classes that served as the genesis for this survey, and to Stacy Carlson, Benjamin Friedman, John Leahy, and Neil Mehrotra for useful comments." en. In: *Handbook of Monetary Economics*. Ed. by Benjamin M. Friedman and Michael Woodford. Vol. 3. Elsevier, Jan. 2010, pp. 183–229 (cit. on p. 28).

[32] N. Gregory Mankiw, Ricardo Reis, and Justin Wolfers. *Disagreement about Inflation Expectations*. Working Paper 9796. Series: Working Paper Series. National Bureau of Economic Research, June 2003. DOI: 10.3386/w9796. URL: https://www.nber.org/papers/w9796 (visited on 11/18/2021) (cit. on p. 5).

[33] Jacob A. Mincer and Victor Zarnowitz. "The Evaluation of Economic Forecasts". In: *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. NBER, 1969, pp. 3–46 (cit. on p. 5).

[34] Silvia Miranda-Agrippino and Hélène Rey. "U.S. Monetary Policy and the Global Financial Cycle". In: *The Review of Economic Studies* 87.6 (Nov. 2020). {\label{1}\Acrobatmenu{GoBack} to Text}, pp. 2754–2776. ISSN: 0034-6527 (cit. on p. 6).

[35] Tanya Molodtsova, Alex Nikolsko-Rzhevskyy, and David H. Papell. “Taylor rules with real-time data: A tale of two countries and one exchange rate”. en. In: *Journal of Monetary Economics*. Contributions to Macroeconomics in Honor of John Taylor 55 (Oct. 2008), S63–S79. ISSN: 0304-3932 (cit. on p. 6).

[36] John F. Muth. “Rational Expectations and the Theory of Price Movements”. en. In: *Econometrica* 29.3 (July 1961), p. 315. ISSN: 00129682 (cit. on p. 4).

[37] Athanasios Orphanides. “Monetary Policy Rules Based on Real-Time Data”. en. In: *MONETARY POLICY RULES* 91.4 (2001), p. 56 (cit. on p. 6).

[38] Christopher Otrok and Charles H. Whiteman. “Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa”. In: *International Economic Review* 39.4 (1998). Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University], pp. 997–1014. ISSN: 0020-6598 (cit. on pp. 22, 23).

[39] Andrew J. Patton and Allan Timmermann. “Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion”. en. In: *Journal of Monetary Economics* 57.7 (Oct. 2010), pp. 803–820. ISSN: 0304-3932 (cit. on pp. 5, 7).

[40] Thomas Sargent and C.A. Sims. “Business Cycle Modeling Without Pretending to Have Too Much a Priori Economic Theory”. In: *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis, Oct. 1977, pp. 45–109 (cit. on p. 22).

[41] Steven M. Sheffrin. *Rational Expectations*. en. Cambridge University Press, June 1996. ISBN: 978-0-521-47939-4 (cit. on p. 5).

[42] Christopher A. Sims. “Implications of rational inattention”. en. In: *Journal of Monetary Economics*. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information 50.3 (Apr. 2003), pp. 665–690. ISSN: 0304-3932 (cit. on pp. 25, 28).

[43] Nicholas S. Souleles. “Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys”. en. In: *Journal of Money, Credit and Banking* 36.1 (2004). Publisher: Blackwell Publishing, pp. 39–72 (cit. on p. 5).

[44] James Stock and Mark Watson. *Business Cycles, Indicators and Forecasting*. Pages: 95-156. University of Chicago Press, 1993 (cit. on pp. 3, 22).

[45] James H. Stock and Mark W. Watson. *Implications of Dynamic Factor Models for VAR Analysis*. Working Paper. July 2005 (cit. on p. 22).

[46] James H. Stock and Mark W. Watson. “New Indexes of Coincident and Leading Economic Indicators”. In: *NBER Macroeconomics Annual 1989, Volume 4*. MIT Press, 1989, pp. 351–409 (cit. on pp. 3, 21, 22).

[47] Michael Woodford. “Imperfect Common Knowledge and the Effects of Monetary Policy”. In: (Jan. 2002) (cit. on p. 25).

Appendix A Proof of Identification of the Dynamic Factor Model

The objective of this mathematical proof is to show that the two identification restrictions outlined in section 4.1.1 uniquely identify the factor loadings and dynamic factors outlined in the dynamic factor model specified in equations (5) through (7).

In a similar approach as Bai and Wang (2015) [5], let C be a full-rank $m \times m$ rotation matrix. Left-multiply the dynamic factor from equation (7) by C and right-multiply the matrix of unknown factor loadings, λ, by C^{-1}. This rotation defines the new factor $\tilde{f}_t = Cf_t$ which allows us to restate equation (7) as,

$$\tilde{f}_t = \Phi_1^f \tilde{f}_t + ... + \Phi_q^f \tilde{f}_{t-q} + C\eta_t^f. \tag{11}$$

After the rotation, the measurement equation (5) becomes,

$$e_t = \lambda C^{-1} \tilde{f}_t + u_t. \tag{12}$$

To establish identification, it must be that the only C that should be acceptable is a diagonal matrix with either 1 or -1 on the diagonal, in which case, both the factors and factor loadings are identified up to a sign change.

From equation (9) the first normalization requirement implies that $var\ (C\eta_t^f) = I_m$. This implies that,

$$CC' = I_m, \tag{13}$$

or that C is an orthogonal matrix.[15]

To address the second normalization requirement that λC^{-1} be a lower triangular matrix with ones on the diagonal, let

$$\lambda = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \lambda_{m1} & \dots & \dots & 1 \\ \vdots & \dots & \dots & \vdots \\ \lambda_{n1} & \dots & \dots & \lambda_{nm} \end{bmatrix} \text{ and } C^{-1} = \begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mm} \end{bmatrix}.$$

[15]The implication here is also that $CC^{-1} = I_m$ and therefore, $C' = C^{-1}$

Carrying out the operation λC^{-1} gives us,

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \lambda_{m1} & \dots & \dots & 1 \\ \vdots & \dots & \dots & \vdots \\ \lambda_{n1} & \dots & \dots & \lambda_{nm} \end{bmatrix} \begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mm} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21}^{*} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \lambda_{m1}^{*} & \dots & \dots & 1 \\ \vdots & \dots & \dots & \vdots \\ \lambda_{n1}^{*} & \dots & \dots & \lambda_{nm}^{*} \end{bmatrix}. \tag{14}$$

Equation (14) shows that the $n \times m$ matrix, λC^{-1}, is a lower triangular matrix where $\lambda_{ij}^{*} = 0$ $\forall\ j > i$ and assuming that $\lambda_{ii}^{*} \neq 0$, $i = 1, ..., m$. A further implication of equation (13) is that $C^{-1}(C^{-1})' = I_m$. Hence, multiplying both sides of equation (14) by $(C^{-1})'$ we have,

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \lambda_{m1} & \dots & \dots & 1 \\ \vdots & \dots & \dots & \vdots \\ \lambda_{n1} & \dots & \dots & \lambda_{nm} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21}^{*} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \lambda_{m1}^{*} & \dots & \dots & 1 \\ \vdots & \dots & \dots & \vdots \\ \lambda_{N1}^{*} & \dots & \dots & \lambda_{nm}^{*} \end{bmatrix} \begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mm} \end{bmatrix}, \tag{15}$$

from which we can deduce that $c_{ij} = 0$ $\forall i, j$ when $i > j$. This proves therefore that C^{-1} is a diagonal $m \times m$ matrix. Further, since $C^{-1}(C^{-1})' = I_m$, it must be that $c_{ii} = 1$ for all $i = 1, ..., m$, so that the rotation matrix C is also a diagonal matrix with either 1 or -1 on the diagonal. This proves, therefore, that the dynamic factors and the associated factor loadings are identified up to a sign change.

Finally, note that the normalization assumption that $\lambda_{ii}^{*} = 1$, $i = 1, ..., m$, imposes a positive sign, therefore, both the dynamic factors and the associated factor loadings are fully identified. **Q.E.D**.