How quickly do forecasters incorporate news? Evidence from cross-country surveys

Isiklar, Gultekin; Lahiri, Kajal and Loungani, Prakash

2006

This revision: July 20, 2005

HOW QUICKLY DO FORECASTERS INCORPORATE NEWS?
Evidence from Cross-country Surveys

GULTEKIN ISIKLAR
University at Albany - SUNY

KAJAL LAHIRI*
University at Albany - SUNY

PRAKASH LOUNGANI
International Monetary Fund

SUMMARY

Using forecasts from *Consensus Economics Inc.,* we provide evidence on the efficiency of real GDP growth forecasts by testing if forecast revisions are uncorrelated. As the forecast data used are multi-dimensional—18 countries, 24 monthly forecasts for the current and the following year and 16 target years—the panel estimation takes into account the complex structure of the variance covariance matrix due to propagation of shocks across countries and economic linkages among them. Efficiency is rejected for all 18 countries: forecast revisions show a high degree of serial correlation. We then develop a framework for characterizing the nature of the inefficiency in forecasts. For a smaller set of countries, the G-7, we estimate a VAR model on forecast revisions. The degree of inefficiency, as manifested in the serial correlation of forecast revisions, tends to be smaller in forecasts of the US than in forecasts for European countries. Our framework also shows that one of the sources of the inefficiency in a country's forecasts is resistance to utilizing foreign news. Thus the quality of forecasts for many of these countries can be significantly improved if forecasters pay more attention to news originating from outside their respective countries. This is particularly the case for Canadian and French forecasts, which would gain by paying greater attention than they do to news from the United States and Germany respectively.

Keywords: Consensus economics, forecast inefficiency, GMM, VAR, panel data.

JEL classifications: E17, E37, F47.

* Corresponding author: K. Lahiri, Department of Economics, SUNY-Albany, Albany, NY 12222. USA
E-mail:klahiri@albany.edu, Fax: (518) 442-4736, Phone: (518) 442-4758.

1 INTRODUCTION

This study tests forecast efficiency and measures the degree of inefficiency in a cross-country panel of fixed-event GDP forecasts. The tests are based on the property that successive forecasts of the same event should be uncorrelated, a property emphasized by Nordhaus (1987) and explained succinctly as follows:

> "If I can look at your most recent forecasts and accurately say, "Your next forecast will be 2% lower than today's, then you can surely improve your forecast" (Nordhaus, p. 673).

Our paper makes two contributions to the literature on assessment of cross-country forecasts. Past studies have tended to evaluate the forecasts country-by-country without considering the commonality of shocks and their propagation among countries in their samples.[1] Our first contribution is to consider cross-country and inter-temporal correlation of forecast revisions using a Generalized Method of Moments (GMM) framework. We thus provide tests of efficiency in a multi-dimensional panel data setting incorporating correlation of forecast revisions among countries, forecast horizons and target years.

A second contribution of this paper is to develop an econometric framework that can shed light on the nature of the inefficiencies in forecasts, in particular the *degree* and the *source* of inefficiencies and the *speed* with which new information is incorporated into forecasts. Our idea is to represent successive forecast revisions for a set of countries using a VAR model and use the coefficients of the model to estimate the extent to which new information is incorporated in successive fixed-event forecasts. The estimated model tells us the degree of inefficiency by showing how long it takes for correlations among successive forecast revisions to die out. The cross-temporal correlations of a country's forecast revisions with those of other countries provide evidence on sources of inefficiencies. Thus our framework provides answers to questions such as: Does foreign news, as compared to domestic news, take longer to be fully incorporated in the forecast of a specific country? Do forecasts incorporate news emanating from other countries efficiently? We also use the variance decompositions to assess the speed with which new information is incorporated in successive fixed-event forecasts. In short, we take the test of efficiency from being purely normative to one that is operationally useful: it provides answers to questions such as those posed above and can also help improve forecast efficiency. The estimated VAR model also provides some evidence on cross-country GDP linkages that arise through the propagation of shocks across countries or through trade and financial linkages among countries.

The main findings of this paper can be summarized as follows. First, our efficiency tests show that the real GDP growth forecasts of the eighteen industrialized countries in our sample are not efficient. This finding is consistent with that of many previous studies that have used widely different data sets.[2] We also provide additional evidence from tests of efficiency based on the direction of forecast revisions. Again, the preponderance of evidence suggests that the forecasts are inefficient.

Having concluded from these initial tests that forecasts are not efficient, we document the nature and the extent of inefficiency in the G-7 real GDP forecasts using the impulse responses and variance decompositions derived from estimating VAR models on forecast revisions. We estimate a VAR model and use generalized impulse responses to measure the degree of inefficiency in the utilization of individual country shocks. Our second finding is to confirm, in this richer framework, that forecast revisions are serially correlated and to provide evidence on the extent and persistence of the correlation and how these differ across countries. For instance, Germany and the U.K. show much more persistent serial correlation in forecast revisions (i.e. much more inefficiency) than the other countries.

Third, there are also interesting differences in the extent to which a country's forecast revisions are correlated with the forecast revisions of other countries. While Japan is largely insulated from the remaining G-7 countries, the other countries display

---

[1] Studies on the accuracy of cross-country forecasts and tests for rationality (Muth, 1961) include Ash *et al.* (1998), Öller and Barot (2000), Pons (2000), Batchelor (2001), and Loungani (2001).
[2] See Fildes and Stekler (2002) for a recent survey.

dependence on the U.S. and, in many instances, on one another (e.g. France on Germany). This suggests that countries could improve the efficiency of their forecasts through greater incorporation of information from the other countries.

Fourth, we use cumulative "intertemporal variance decompositions" of forecast revisions as a way of summarizing information on the *speed* with which new information is incorporated into forecasts. We find that 90% of new information is used most promptly in the US (within 2 months), followed by Japan, Canada and Italy (3 months), the UK and France (4 months), and Germany (5 months).

In the next section we introduce the data set. In section 3, we present the model, the methodology for conducting GMM tests of efficiency, and the structure of the variance-covariance matrix of revisions. In section 4, we present several characteristics of the forecast revision data and results of the GMM tests for efficiency. Section 5 provides the VAR model framework, the empirical results related to the nature of efficiency, and the so-called intertemporal variance decomposition curve. In the last section we summarize the results.

2 CONSENSUS FORECASTS

Since October 1989, the *Consensus Economics Inc.* service has been polling—initially a few but currently more than 600—forecasters each month and recording their forecasts for principal macroeconomic variables (including GDP growth, inflation, interest rates and exchange rates) for a set of countries. Forecasts are made for the current year (based on partial information about developments in that year) and for the following year. The number of panelists ranges from 10 to 30 for each of the countries, and for the major industrialized countries the panelists are generally based in countries they forecast. For example, for the G-7 countries, most *Consensus Economics* panelists are based in the country whose macroeconomic variables they forecast.

We study the consensus forecasts (i.e., averages of the individual responses) of annual average real GDP growth. Survey respondents make their first forecasts when there are 24 months to the end of year; that is, they start forecasting GDP for the target year in January of the previous year, and their last forecast is reported in the beginning of December of the target year. So for each country and for each target year we have 24 forecasts of varying horizons. Our data set ranges from October 1989 to June 2004. The countries we study are the 18 industrialized countries for which forecasts are available from *Consensus Economics*[3] and some of the results are for a subset of that group, the G-7 countries.

Only a handful of studies have used the *Consensus Forecasts* data set. These include Artis and Zhang (1997), Batchelor (2001), Harvey *et al*. (2001), Loungani (2001) and Gallo *et al.* (2002). Among these studies Harvey *et al.* (2001) and Loungani (2001) contain formal tests of forecast efficiency for, respectively, the UK and a large set of countries; but neither paper made allowance for the cross-country correlation of shocks, a shortcoming that is addressed in this paper.

3 TESTING FORECAST EFFICIENCY AND STRUCTURE OF CROSS-COUNTRY FORECASTS

3.1 Efficiency Tests

Our data set is composed of forecasts of a fixed event (year-over-year average real GDP growth) made monthly before the end of the target year, i.e., made at different forecast horizons. As Clements and Hendry (1998) note, the common way of testing efficiency using fixed event forecasts is still essentially the one suggested by Nordhaus (1987). Nordhaus defines a notion of efficiency that can be tested using forecast revisions. Strong efficiency requires that all available information, including appropriate knowledge about the structure of the economy is incorporated in the forecast. This argument can be expressed by the condition $E[r_{i,t,h} \mid \Phi_{i,t,h+1}] = 0$, where the forecast

[3] Forecasts for some additional industrialized countries are reported in *Asia-Pacific Consensus Forecasts*, an offshoot of the original publication.

revision $r_{i,t,h} = f_{i,t,h} - f_{i,t,h+1}$ is the difference between two successive forecasts in country i for the same target year t, and $\Phi_{i,t,h+1}$ is the information set used by forecasters in country i when the forecast horizon is h+1. Since the information requirement in strongly efficient forecasts is almost impossible to specify, it has been very difficult to implement tests for strong efficiency in practice. Because of this practical limitation in testing efficiency using the unobserved $\Phi_{i,t,h+1}$, Nordhaus proposed a new concept called 'weak efficiency' in which the set of all past forecasts is substituted for $\Phi_{i,t,h+1}$. Since past revisions are clearly in the information set of the forecasters, efficient forecasts require that the revisions should be uncorrelated with their past values. So a necessary condition for testing weak efficiency for fixed-event forecasts is $E\left[r_{i,t,h} \mid r_{i,t,h+1}, r_{i,t,h+2}, ..., r_{i,t,h+H}\right] = 0$, where $r_{i,t,h+k}$ is the forecast revision for the target date t and when the forecast horizon is $h+k$.

To test for weak efficiency, the common practice is to see if $\beta_1 = 0$ in the following regression: $r_{i,t,h} = \beta_1 r_{i,t,h+k} + u_{i,t,h}$, where i denotes the country, t the target year, h the forecast horizon and $k \geq 1$. If β_1 is found to be significantly different from zero, the conclusion is that forecasters are not efficient, that is, they do not update their forecasts to fully reflect the new information that has arrived since some previous revision of their forecast. If β_1 is found to be significantly greater than zero but less than one, it implies that the forecasters adjust the forecasts slowly: they smooth their forecasts. This type of forecast efficiency tests has been employed in several studies. See, for example, Batchelor and Dua (1991), Davies and Lahiri (1995, 1999), Clements (1995, 1997) and Harvey *et al.* (2001).

There are several advantages of using this test of forecast efficiency. As Nordhaus (1987) demonstrates, in certain circumstances this test may have more power in detecting inefficiency over tests that use rolling event forecasts, i.e. when target year t changes and forecast horizon h is fixed. The additional power of the test comes from its ability to detect smoothing of forecasts following the arrival of new information. Especially when the number of fixed targets is small, as in our case, tests based on rolling-event forecasts are less powerful in detecting forecast smoothing. Another advantage of the current test is that it uses only forecast revisions and not actual data. Hence it sidesteps the contentious issue of which version of the 'actual' data one should use, a preliminary release of the data or later revisions.

We also provide evidence on efficiency from a related non-parametric test. If forecasts are efficient then an outsider should not be able to predict the direction of the next forecast revision. That is, the direction of this month's forecast revision should not have any predictive value for future forecast revisions. Thus one can document the persistence in the direction of forecast revisions using contingency tables and then test for the independence of current and future forecast revisions.[4] In addition to providing further evidence on efficiency, this second test is useful in overcoming a potential weakness of the first test. It is now well known that over part of our sample period, forecasters were surprised by the 'new economy' phenomenon. There may have been a pattern of positive serial correlation in output shocks over this period that was a departure from historical norms, thus causing growth forecasts to be revised up steadily. If that is the case, then the evidence from the first test could mistakenly attribute some of the resulting serial correlation in forecasts revisions to a lack of efficiency.[5] The second test investigates this possibility. If the directional persistence in forecast revisions comes predominantly from *positive* revisions, this would lend some credence to the view that an abnormal pattern of positive serial correlation in output shocks, rather than inefficiency, may be the cause of serially correlated forecast revisions.

3.2 The Panel Data Model

Under the null hypothesis of efficiency, the error $u_{i,t,h}$ will have a special structure. Davies and Lahiri (1995) provide an econometric framework for analyzing fixed event forecast errors in a multi-dimensional panel data setting that can be used to specify the

[4] These types of cross-tables are frequently used in evaluating the accuracy of forecasts by comparing the direction of the change in forecasts with the actual direction of change. See Merton (1981), Henriksson and Merton (1981), and Pesaran and Timmerman (1992). Ash *et al.* (1998) contains additional details on these tests.

[5] Note that if the shocks are serially correlated then a forecaster who uses full information should incorporate this serial correlation in his/her forecasts. However, if the serial correlation cannot be observed in real time for some reason, then the observed persistence in revisions may be consistent with efficiency.

covariance structure of $u_{i,t,h}$. Davies and Lahiri (1995) test for forecast rationality, using data from the Blue Chip Survey of Professional Forecasters, when the forecast errors are correlated across forecasters, target years and forecast horizons. The main difference between the structure of their data set from the one used here is that in their data set there are multiple forecasts (one by each forecaster) for the same target variable, whereas here we have a single forecast (the consensus forecast) for the target variable. Another difference stems from the structure of the correlation in errors. In the Davies and Lahiri (1995) setting, the forecasts are correlated only because of common aggregate shocks, whereas here the forecasts of two countries can be correlated not only because of common aggregate shocks but also the economic interdependence between them. Since our efficiency tests use forecast revisions, most of the attention will be on the structure of the forecast revisions, $r_{i,t,h} = f_{i,t,h} - f_{i,t,h+1}$.

Since the forecasts span a twenty-four month period with monthly revisions, the structure of the variance-covariance matrix must accommodate correlations of: (i) contemporaneous forecast revisions within and across countries for the same target year; (ii) contemporaneous revisions for the same country but for different target years; and (iii) contemporaneous revisions across countries, but for different targets. In addition, the estimation has to account for the potential aggregation bias that may arise from averaging individual forecasts to generate the consensus forecasts.

In order to model the transmission of shocks across countries, we assume that each country receives an idiosyncratic shock ($\varepsilon_{i,t,h}$) that gets transmitted to other countries either instantaneously or with lags depending on the economic linkages between the countries. The well-documented co-movements of many economies can be accounted for by the presence of interdependencies via goods or asset markets, which transmit country-specific shocks across national boundaries causing contemporaneous forecast revisions to be correlated across countries.

Next, since forecasts are made each month for both the current and the next year, the shock in one month that has an effect on the current year real GDP growth will be correlated with the shock in the same month having an effect on next year's GDP growth, that is, $Cov\left(r_{i,t,h}, r_{j,t+1,h+12}\right) \neq 0$ for every h < 12 and for all *(i,j)*.

Finally, our results may be influenced by possible aggregation bias because of the use of the mean (i.e., consensus) forecasts instead of the individual forecasts. Suppose that there is a large discrepancy in the timing of when the individual forecasts are made or there are frictions affecting the speed with which news gets to different forecasters. Davies and Lahiri (1995) show that aggregation can result in a significant first order correlation coefficient in the forecast revisions of the mean (consensus) forecast. To avoid mistaking serial correlation resulting from aggregation bias for inefficiency, we regress the current revisions on the second or third lagged revisions and include an MA(1) term in the variance covariance matrix of forecast revisions.

In short, the various non-zero components of the covariance matrix of forecast revisions we estimate for the GMM estimation are[6]:

1) Revision variance for country i at forecast horizon h: $Var\left(r_{i,t,h}\right) = \sigma_i^2$ for h = 1,...,23

2) Own country covariance for consecutive targets: $Cov\left(r_{i,t,h}, r_{i,t+1,h+12}\right) = \omega_i$ for h < 12

3) Cross country covariance for the same target: $Cov\left(r_{i,t,h}, r_{j,t,h}\right) = \gamma_{ij}$ for h = 1,...,23

4) Cross country covariance for consecutive targets: $Cov\left(r_{i,t,h} . r_{j,t+1,h+12}\right) = s_{ij}$ for h < 12

5) Own country covariance due to aggregation: $Cov\left(r_{i,t,h}, r_{i,t,h+1}\right) = m_i$ for h = 1,...,22.

Computational details on the estimation of this variance-covariance matrix are given in the Appendix.

[6] The variation in the number of respondents over time and across countries in the computation of consensus forecasts could introduce additional degree of uncertainty in our estimates.

4 TESTS OF EFFICIENCY: EMPIRICAL RESULTS

To show the importance of cross-country spillovers and commonality of news, we present the correlation of the forecast revisions across our 18 sample countries in Table I. These correspond to the elements γ_{ik} defined above. As seen in this table, most of these cross-country correlations are large; moreover, forecast revisions for countries with close links are highly correlated. For example, the correlation of forecast revisions is 0.52 between USA and Canada and 0.52 between France and Germany. In contrast, the correlations between Japan and most of the countries are smaller than 0.25. These contemporaneous correlations can be due to both production and consumption interdependencies between countries or due to common exogenous shocks even in the absence of such interdependencies.[7,8] In short, these calculations show that the variance covariance matrix of forecast revisions has significant off-diagonal elements, and the covariance structure to be used in GMM estimation is consistent with the data.

We test for efficiency of GDP growth forecasts using the two different tests discussed in section 3.1. Initially, we perform GMM estimation and show that the forecast revisions are serially correlated. Then, we show that the *direction* of revisions in GDP growth forecasts can be predicted by an outsider, also suggesting inefficient use of information.

To test for efficiency, we estimate the regression: $r_{i,t,h} = \beta_1 r_{i,t,h+2} + u_{i,t,h}$. Note that the regressor is the second lagged forecast revision.[9] Table II provides estimates of β_1 (the OLS and GMM estimates are identical for exactly identified systems) along with the t-statistics for the OLS and GMM estimates. When the data for all countries are pooled, β_1 is estimated as 0.33, and is highly significant under both the OLS and GMM

---

[7] The Pesaran (2004) test statistic for cross section dependence is found as 69.03, which is significant at the 1% significance level. The LM test of Breusch and Pagan (1980) is also significant at the 1% level.

[8] We also calculated the correlations in forecast revisions that are reported at the same time but for two successive target years. These correspond to the own country covariances and the cross-country covariances. These forecasts are indeed highly correlated. Also, the closely related countries have larger correlations than others and, as expected, the own country correlation is always the largest.

[9] Using the first lagged value of the forecast revision as the regressor in the efficiency tests gave results that were very similar to those reported in Table II. However, as discussed, we use the second lagged revision to avoid possible aggregation bias.

estimation. The same regression is also estimated country-by-country using the relevant portion of the variance-covariance matrix. In all cases the estimates of β_1 are found to be statistically significant at the 5% level of significance. Thus, we conclude that the forecasts are not efficient.

Table III presents summary measures of the frequency distribution of the direction of forecast revisions and the expected frequencies under the null hypothesis of independence. For brevity, we only present the important diagonal elements of the full 3x3 contingency table. The first column of numbers shows the frequency of negative revisions that were preceded three months earlier by negative revisions (the 'negative/negative' cell) and compares that with the expected frequency; likewise, the second column shows the incidence of positive revisions preceded by positive revisions (the 'positive/positive' cell) and the expected frequency in that case.[10] For example, we find that for Canadian output growth forecasts, the number of forecast revisions that are positive both now and three months ago is 84, while the expected value of the positive/positive cell is 54.9 under the null of independence. Similarly, the frequency of forecast revisions that are negative both now and three months ago is 92, while the expected value of the negative/negative cell is 70.8 under independence.

It is evident that for all of the entries in the table, observed frequencies are larger than the expected frequencies. Moreover, the resulting chi-square statistic is significant. We thus conclude that there is very strong directional persistence of forecast revisions among G-7 countries; it is easy to predict the direction of future forecast revisions, even those that are three months in the future.[11]

If forecasters were surprised by the exceptional growth of the new economy in the 1990s, one would have expected the 'positive/positive' cell to contribute far more to the chi-square statistic than the 'negative/negative' cell. This is not the case. As it can be inferred from Table III, the contribution of the 'negative/negative' cell, although usually

[10] For the sake of brevity, we do not provide the full 3x3 cross-table of past negative, zero, and positive revisions against current negative, zero and positive revisions.

[11] As expected, the support for inefficiency increases when we use 1 month ahead or 2 months ahead forecast revisions instead of the 3 months ahead revisions used in the results reported in Table III. We also find support for inefficiency for the non-G-7 countries in our sample.

a bit smaller than the contributions of the 'positive/positive' cell, is quite significant. Thus, we conclude that the source of the observed inefficiency in forecasts is not the result of an abnormal pattern of positively correlated shocks over the 1990s but is due to the under- utilization of new information.

5 MEASURING THE DEGREE OF INEFFICIENCY

The inefficiency documented in the tests above implies that some part of new information is not incorporated in the forecasts immediately. We develop measures of inefficiency that indicate *how much* of the new information is incorporated immediately and how much is incorporated over time.

5.1 The VAR Model

Our approach is simple and is based on the tests of forecast efficiency. If forecasts are efficient, forecast revisions represent the effect of the new information on the target variable. This has implications for the moving-average (MA) representation of the forecast revisions. If the forecasts are efficient, forecast revisions will follow an MA(0) process ($r_{i,t,h} = \varepsilon_{i,t,h}$), and there will be no correlation between two successive forecast revisions. But if forecasts are not efficient, forecast revisions will be correlated. Moreover, there could be correlation not only with own-country forecast revisions but with revisions in the forecasts of other countries. So one could envisage a general VAR(p) model of forecast revisions:

$$r_{t,h} = c + B_1 r_{t,h+1} + B_2 r_{t,h+2} + \ldots + B_p r_{t,h+p} + \varepsilon_{t,h} \qquad (1)$$

where $r_{t,h}$ denotes a (n $\times$ 1) vector containing the forecast revisions of the n countries when the forecast horizon is h and target year is t, B_k denotes the (n $\times$ n) matrix of coefficients of $r_{t,h+k}$, and p is the chosen lag length.[12] *VAR(p)* can be rewritten in *VMA(∞)* form as

$$r_{t,h} = \mu + M_0 \varepsilon_{t,h} + M_1 \varepsilon_{t,h+1} + M_2 \varepsilon_{t,h+2} + \cdots. \tag{2}$$

where M_k satisfies $M_k = B_1 M_{k-1} + B_2 M_{k-2} + \cdots + B_p M_{k-p}$ with $M_0=I$ and $M_k=0$ for k < 0. A VAR(3) is estimated with the G-7 countries – the U.S., Japan, Germany, the U.K., France, Italy, and Canada. The optimum lag length is chosen by AIC.

As mentioned earlier, if forecasts incorporate all the available information efficiently, forecast revisions should be uncorrelated with their lagged values as well as the lagged values of forecast revisions of other countries. The estimated VAR system presents us with an important tool to understand the dynamics of the forecasting process in more detail than the simple correlations. In its usual interpretation, impulse responses trace the effect of a one standard deviation shock to one of the innovations on the future values of other variables in the system. Our variables are revisions of real GDP growth forecasts of the sample countries; hence impulse responses show the responses of forecast revisions to innovations over time. But under perfect efficiency, forecast revisions should respond fully to the shocks immediately. If the forecast revisions do not respond to the shocks immediately, i.e. if there are nonzero impulse response values when horizon is greater than 1, then it means that forecasts are not efficiently using the information immediately, and some of the information is being utilized in the later forecast revisions. In other words, impulse responses of the forecast revisions show the dynamics of how shocks are absorbed in the forecast revisions over time. The longer it takes for the responses to go to zero, the greater is the degree of forecast inefficiency.

The orthogonalized impulse responses and the associated variance decomposition are sensitive to the ordering of the countries in the VAR. Because of this, we used generalized impulse responses and variance decompositions which are ordering-free. The generalized impulse responses were introduced by Koop, Pesaran and Potter (1996) for nonlinear systems. Pesaran and Shin (1998) proposed the method for an ordering free

[12] Panel VAR models have been used to analyze transmission of business cycles in multi-country models

solution in the VAR analysis, and they show that $n \times 1$ vector of k period ahead generalized impulse response of the effect of a one-standard deviation shock in the *j*-th country forecast revision equation is given by

$$\psi_j(k) = \sigma_{jj}^{-1/2} M_k \Omega e_j \tag{3}$$

where e_j is the *j*-th column of an identity matrix and $\Omega = E(\varepsilon_{t,h}\varepsilon'_{t,h}) = \{\sigma_{ij}, i, j = 1, 2, ..., n\}$. M_k have been defined before. Note that, Ω has a sample estimate of $\hat{\Omega} = \left(1/TH\right)\Sigma_t \Sigma_h \hat{\varepsilon}_{t,h}\hat{\varepsilon}'_{t,h}$ where $\hat{\varepsilon}_{t,h}$ is (7×1) residual vector from the *VAR(3)* model.

Figures 1 to 7 show the estimated generalized impulse responses (bold line) with their 2 SE bands (dashed lines). First, consider the impulse responses to own-country shocks in each of the Figures 1 - 7. There is ample evidence of inefficiency: In all seven cases, the evidence shows that revisions do not respond to the own-country shocks immediately. A closer look reveals the extent of the inefficiency varies across countries. Consider, for instance, the difference between the U.S. and German cases. In Figure 1, the panel labeled “Response of US to US” shows that the impulse responses are not statistically significant from zero after month 4. In contrast, in the panel labeled “Response of GE to GE” in Figure 3, the responses remain significant until month 9. In terms of utilizing own-country information the evidence shows that the U.K. and Germany are the most inefficient countries.

Second, consider the ‘off-diagonal’ elements, the panels that show the responses of the forecast revision of one country to the forecast revisions in other countries. For most of the countries, forecast revisions do show a significant dependence on the shocks of other countries. Two exceptions are Japan and the UK. All countries show strong and persistent responses to US shocks implying that forecast efficiency could be improved if forecasters were to show greater awareness of news emanating from the United States. In addition, Germany shows persistent response to France and the UK, France shows persistent response to Germany, Italy shows persistent response to France and the UK, and Canada shows persistent response to the UK.

---

using macro data (*cf*., Canova and Marrinan (1998), Stock and Watson (2003), and Pesaran *et al*. (2003)).

What is the relative importance of own-country shocks and cross-country shocks in forecast revisions? To answer this, in Table IV, we present the generalized forecast error variance decompositions at the steady state.[13] It is clear from Table IV that GDP forecasts of most of the countries are strongly influenced by foreign country information in our sample. Canadian and French forecast revisions have the largest off-diagonal contributions. For example, US shocks account for 36 percent of the variation in Canadian GDP forecast revisions and German shocks account for 27 percent of the French GDP forecast revisions. The dependence on foreign information is also prevalent for Italy, Germany, and the US, but to a lesser extent. Of the G-7 countries, Japan seems to be most immune to foreign shocks.

5.2 Intertemporal Variance Decompositions - Aggregate measure of inefficiency

The variance decompositions presented in Table IV above provide estimates of the relative importance of domestic *vis-a-vis* foreign shocks in explaining forecast revision variance in the long run. But another important feature of forecasts is the *speed* of forecasters' response to aggregate news over time. To compute this, we need to see how much of the variation in forecast revisions is accounted for by current innovations and how much of it is accounted for by past innovations. Thus, we decompose the variation in forecast revisions over time into new and old components using cumulative 'intertemporal variance decompositions'.

To motivate the use of this concept, consider the following example. Suppose not all of the information of the current period is being used fully, some of it is used one period later and the remainder two periods later. Then we would expect to observe an MA(2) model for the forecast revisions: $r_{t,h} = \varepsilon_{t,h} + m_1\varepsilon_{t,h+1} + m_2\varepsilon_{t,h+2}$. In this example, $\varepsilon_{t,h+1}$ is the news that became available last period and m_1 is the coefficient that represents the usage of the news now. In this case, the percentage of the variation due to immediate usage of the new information is

[13] Notice that, in general, generalized variance decompositions do not add up to 100 percent due to non-zero covariances between the original country shocks, see Pesaran and Shin (1998).

$Var\left(\varepsilon_{i,t,h}\right) \Big/ Var\left(\varepsilon_{i,t,h} + m_1 \varepsilon_{i,t,h+1} + m_2 \varepsilon_{i,t,h+2}\right) = 1 \big/ (1 + m_1^2 + m_2^2)$ and the cumulative percentage of the variation within one month is $(1 + m_1^2) \Big/ \left(1 + m_1^2 + m_2^2\right)$.

Generalizing this example, we can decompose the forecast revisions variance using equation (2). For country *k*, the percentage of revision variation due to the immediate usage of the current information is

$$\theta_{k,0} = \frac{e_k' M_0 \Omega M_0{}' e_k}{\sum_{i=0}^{\infty} e_k' M_i \Omega M_i{}' e_k} \tag{4}$$

where e_k is the k-th vector of the identity matrix. The numerator of equation (4) is the i-th diagonal element of the total forecast error variance at horizon zero and $\sum_{i=0}^{\infty} e_i{}' M_i \Omega M_i{}' e_i$ is nothing but the variance of *k-th* element of $r_{t,h}$. Hence $\theta_{k,0}$ gives the percentage of the variation in revisions accounted for by contemporaneous innovations. Similarly the cumulative percentage of the variation of the revisions within *m- periods* is

$$\theta_{k,m} = \frac{\sum_{i=0}^{m} e_k' M_i \Omega M_i{}' e_k}{\sum_{i=0}^{\infty} e_k' M_i \Omega M_i{}' e_k}. \tag{5}$$

In Table V we present the estimated intertemporal variance decompositions with respect to months elapsed. For example for US, 73% of forecast revision variance is accounted for by current news, and an additional 14% is explained by shocks that occurred one month ago, making the cumulative variance explained to be 87%. In this way, within 2 months more than 90% of the US revision variance is explained. Similar to what we found before, the forecasts of other countries look conspicuously less efficient than the US forecasts. Table V shows Germany to be the least efficient where immediate shocks account for only 54% of the total revision variance. If we set 90% as the threshold level for the information usage, we observe that while US forecasts need two months to reach the threshold, Canadian, Japanese and Italian forecasts need three months, the UK and French forecasts need four months, and German forecasts need 5 months to reach the threshold.

5.3 Comparison with Studies of Forecast Accuracy

A large part of the forecast evaluation literature focuses on accuracy, using statistics such as mean errors, root mean squares (RMSE), and mean absolute errors (MAE). While inefficient use of available information, which is the focus of this paper, is an important driver of forecast errors, it is not the only one. Forecast errors are also composed of information that was not available to forecaster at the time of forecast such as data revisions, or structural changes affecting the data generating process after the forecasts were produced. Nevertheless, it is possible for the two approaches to give somewhat similar results when the two information sets are similar, that is, when, *ex post*, observed structural changes or shocks do not play a major role in the process.

A comparison of our results with those in the literature on forecast accuracy does indeed suggest some similarity in the findings. Recall that one of the main results from our VAR analysis is that during 1990-2001 the real GDP growth forecasts for Germany and France incorporate new information at a speed slower than those for US, Japan and Canada. There are several studies that compare the performance of real GDP growth forecasts across countries (e.g., Ash *et al*. (1998), Öller and Barot (2000) and Pons (2000)). Since sample periods, countries or forecasting agencies do not match exactly, an exact comparison with other studies is not possible. Ash *et al.* (1998) compare the directional accuracy of, among other target variables, the real GDP growth forecasts of OECD countries and find that, when the forecast horizon is one year, US is the only country whose GDP growth forecasts have "value". They also find that, when the forecast horizon is 6 months, France is the only country whose GDP growth forecasts do not have any value.

Pons (2000) is most comparable to ours. He studies the accuracy of IMF and OECD forecasts for G-7 countries using the sample period of 1971-1995. Pons (2000) finds that during 1983-1995 the current-year IMF and OECD forecasts have the highest accuracy for US and Japan and the lowest accuracy for Germany, France and UK.

Similarly the year-ahead forecasts have the highest accuracy for US and Japan and the lowest accuracy for Germany, UK and France.

Öller and Barot (2000) study the OECD forecasts of only the European countries so an exact comparison is not possible because they do not include US, Canada and Japan. They find that, during 1985-1997, the accuracy of GDP growth forecasts for Germany and France is worse than those for Italy and UK, which again is somewhat consistent with our findings.

There can be many plausible explanations for the observed difference between the US and some European countries in terms of the speed of information usage. One may be related to the structural changes that the European economies have been going through in the 1990s. In addition, the unification of Germany might have created additional uncertainty about the underlying data generating process, which can lead to stickiness in information usage. Another explanation may come not from the inefficiency of the forecasters but that of the statistical agencies processing the available information to produce GDP figures. If the data used by the forecasters are revised in a serially correlated fashion, the forecast revisions will be correlated even if the forecasters are using all the available information. Faust, Rogers and Wright (2005) found that data revisions produced by the statistical agencies of UK, Italy, and Japan are highly predictable; they are less so for the US. This implies that some part of the observed forecaster inefficiency may be due to the inefficiency of the statistical agencies. For all these reasons, the relative stickiness in some of the European forecasts may not necessarily mean that these forecasters are less skillful than their American counterparts.

6 CONCLUSIONS

This study evaluates the performance of cross-country survey forecasts provided by the *Consensus Economics* service and tests the rational expectations hypothesis. If forecasts are generated by rational agents who use all available information, as assumed by the rational expectations hypothesis, then forecasts should be unbiased and efficient.

Using a GMM estimation framework, we found that these forecasts are not efficient. Our tests took into account the complex structure of the variance covariance matrix of forecast revisions due to the propagation of shocks among countries. We rejected the hypothesis of efficiency in each of the 18 countries individually, and also for the pooled data. Based on a well-known chi-square test of independence in the direction of successive forecast revisions, we also reached the same conclusion.

Secondly, we presented a methodology for measuring the degree of inefficiency. If forecasts are inefficient, it means that forecasters do not use all of the new information when they make revisions in their forecasts. We thus decomposed the current forecast revision into 'current' and 'old' news components and measured the variation due to new and old information. Using a VAR framework and intertemporal variance decompositions, we measured the percentage of new information that is being incorporated in the G-7 GDP forecast revisions within a certain period of time. The analysis revealed that the degree of inefficiency is different for different countries, and that forecasters tend to take 2 to 5 months to incorporate 90% of the new information. The degree of inefficiency is least for US, Japan and Canada where more than 90% of the available information is used in less than 4 months. But it takes 4 months for France and the UK and 5 months for Germany forecasts to incorporate 90% of the new information.

Using generalized impulse responses, we also found that news coming from foreign countries take longer to be reflected in the forecasts. Especially, for Canada and France, where foreign shocks seem to play a more prominent role, the forecast accuracy could be significantly increased if more attention is paid to news originating from US and Germany.

Appendix A

The elements of $Cov(r_{i,t,h}, r_{k,t',h'})$ may be directly calculated using the revision series and after averaging out the relevant elements of the revision matrix RR' where R is the vector of forecast revisions stacked in country, target year and forecast horizon respectively.

When the second lagged term is used as the RHS variable in tests of efficiency the number of total observations is 301 for each country. So the dimension of variance covariance matrix is 5418 $\times$ 5418, (with 18 countries 301*18=5418). The number of estimated items of the variance-covariance matrix is as follows:

- The number of diagonal elements (σ_i^2) = 18;
- The number of own country covariance for the same target (ω_i^2) = 18;
- The number of cross-country covariances for the same target (γ_{ij}) =(18 * 17)/2=153;
- The number of cross-country covariances for consecutive targets (s_{ij})=(18*17)/2=153.

All the estimates are computed as the averages of the sample covariances as follows:

$$\hat{\sigma}_i^2 = \frac{1}{TH}\sum_{k=1}^{H}\sum_{t=1}^{T}\tilde{r}_{itk}^2$$

$$\hat{\omega}_i = \frac{1}{\hat{H}(T-1)}\sum_{k=1}^{\tilde{H}}\sum_{t=1}^{T-1}\tilde{r}_{itk}\tilde{r}_{it+1,k+12}$$

$$\gamma_{ij} = \frac{1}{TH}\sum_{k=1}^{H}\sum_{t=1}^{T}\tilde{r}_{itk}\tilde{r}_{jtk}$$

$$\hat{s}_{ij} = \frac{1}{2\hat{H}(T-1)}\sum_{k=1}^{\tilde{H}}\sum_{t=1}^{T-1}\tilde{r}_{itk}\tilde{r}_{jt+1,k+12} + \tilde{r}_{jtk}\tilde{r}_{it+1,k+12} \,,$$

where ~ denotes deviations from averages and $\hat{H}$ is 9 when the second lagged revisions are used as the explanatory variable in equation $r_{i,t,h} = \beta_1 r_{i,t,h+2} + u_{i,t,h}$.

ACKNOWLEDGEMENTS

We thank Editor M. Hashem Pesaran, Roy Batchelor, David Harvey, Lars-Erik Oller, John Robertson, Herman Stekler, Ellis Tallman, and two anonymous referees of the journal for their comments and suggestions on earlier versions. The paper has also benefited from presentations at the 24th International Symposium on Forecasting, Sydney (Australia), the Federal Reserve Bank of Atlanta, and the International Monetary Fund.

REFERENCES

Artis MJ, Zhang W. 1997. International business cycles and the ERM: Is there a European business cycle? *International Journal of Finance and Economics,* 2: 1-16.

Ash JCK, Smyth DJ, Heravi SM. 1998. Are OECD forecasts rational and useful? A directional analysis. *International Journal of Forecasting* 14: 381-391.

Batchelor RA. 2001. How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus, *Applied Economics* 33: 225-35.

Batchelor RA, Dua P. 1991. Blue Chip rationality tests, *Journal of Money, Credit and Banking* 23: No: 4.

Breusch, TS, Pagan AR. 1980. The Lagrange Multiplier test and its application to model specifications in econometrics, *Review of Economic Studies*, 47, 239-53.

Canova F, Marrinan J. 1998. Sources and propagation of international output cycles: common shocks and transmission, *Journal of International Economics*, 46: 133-166.

Clements MP. 1995. Rationality and the role of judgment in macroeconomic forecasting, *Economic Journal*, 105: 410-420.

Clements MP. 1997. Evaluating the rationality of fixed-event forecasts, *Journal of Forecasting*, 16: 225-239.

Clements MP, Hendry DE. 1998. *Forecasting Economic Time Series*, Cambridge University Press; Cambridge, UK.

Davies A, Lahiri K. 1995. A new framework for analyzing survey forecasts using three-dimensional panel data, *Journal of Econometrics* 68: 205-227.

Davies A, Lahiri K. 1999. Reexamining the rational expectations hypothesis using panel data on multi-period forecasts," in *Analysis of Panels and Limited Dependent Variable Models* (eds. C. Hsiao *et al*.), Cambridge University Press: Cambridge, 226-254.

Faust J, Rogers JH, Wright JH. 2005. News and noise in G-7 GDP announcements, *Journal of Money, Credit and Banking*, Forthcoming.

Fildes R, Stekler H. 2002. "The state of macroeconomic forecasting," *Journal of Macroeconomics*, 24: 435-468.

Gallo GM, Granger CWJ, Jeon Y. 2002. Copycats and common swings: the impact of the use of the forecasts in information sets, *IMF staff Papers* 49: 4-21.

Harvey DI, Leybourne SJ, Newbold P. 2001. Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal* 4: 37-55.

Henriksson RD, Merton C. 1981. On market timing and investment performance 2: Statistical procedures for evaluating forecasting skills, *Journal of Business* 54: 513-533.

Koop G, Pesaran MH, Potter SM. 1996. Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics* 17: 107-12.

Loungani P. 2001. How accurate are private sector forecasts? Cross- country evidence from consensus forecasts of output growth. International Journal of Forecasting 17: 419-432.

Merton RC. 1981. On market timing and investment performance 1: an equilibrium theory of value for market forecasts," *Journal of Business* 54: 363-406.

Muth JF. 1961. Rational expectations and the theory of price movements. *Econometrica* 29: 315-335.

Nordhaus W.1987. Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics* 69: 667-674.

Öller L-E, Barot B. 2000. The accuracy of European growth and inflation forecasts. *International Journal of Forecasting* 16: 293-315.

Pesaran MH. 2004. General Diagnostic Tests for cross section dependence in panels, mimeo. University of Cambridge, June 2004.

Pesaran MH, Timmerman AG. 1992. A simple nonparametric test of predictive performance. *Journal of Business and Economics Statistics* 10: 461-465.

Pesaran MH, Shin Y. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58: 17-29.

Pesaran, MH, Schuermann T, Weiner SM. 2003. Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconomic Model. *Journal of Business & Economic Statistics*, 22, 129-162.

Pons J. 2000. The accuracy of IMF and OECD forecasts for G7 countries. *Journal of Forecasting* 19: 53-63.

Stock JH, Watson MW. 2003. Understanding changes in international business cycle dynamics. *NBER Working Papers* No. w9859, July. Cambridge, MA.

Table I. Correlations in Forecast Revisions (Sample Size=333).

	Austria	**Belgium**	**Canada**	**Denmark**	**Finland**	**France**	**Germany**	**Netherl.**	**Ireland**	**Italy**	**Japan**	**Norway**	**Poland**	**Spain**	**Sweden**	**Switz.**	**UK**	**US**
Austria	1.00																	
Belgium	0.54	1.00																
Canada	0.15	0.21	1.00															
Denmark	0.41	0.42	0.27	1.00														
Finland	0.43	0.43	0.31	0.34	1.00													
France	0.32	0.41	0.36	0.29	0.39	1.00												
Germany	0.44	0.46	0.35	0.43	0.35	0.52	1.00											
Netherl.	0.37	0.47	0.37	0.33	0.37	0.56	0.56	1.00										
Ireland	0.20	0.29	0.28	0.26	0.40	0.29	0.27	0.41	1.00									
Italy	0.26	0.37	0.35	0.31	0.32	0.42	0.40	0.44	0.25	1.00								
Japan	0.15	0.15	0.24	0.06	0.13	0.22	0.26	0.21	0.14	0.16	1.00							
Norway	0.22	0.27	0.16	0.28	0.22	0.21	0.28	0.22	0.15	0.19	0.09	1.00						
Poland	0.44	0.45	0.28	0.31	0.37	0.40	0.39	0.41	0.25	0.37	0.14	0.03	1.00					
Spain	0.38	0.44	0.32	0.40	0.36	0.49	0.49	0.44	0.30	0.41	0.19	0.23	0.42	1.00				
Sweden	0.28	0.28	0.35	0.29	0.37	0.34	0.38	0.37	0.17	0.35	0.25	0.12	0.28	0.37	1.00			
Switz.	0.34	0.46	0.26	0.28	0.42	0.44	0.32	0.39	0.31	0.43	0.15	0.22	0.24	0.39	0.26	1.00		
UK	0.08	0.16	0.34	0.21	0.25	0.22	0.27	0.26	0.23	0.24	0.14	0.15	0.11	0.33	0.28	0.19	1.00	
US	0.23	0.28	0.52	0.19	0.34	0.40	0.28	0.36	0.23	0.37	0.22	0.09	0.33	0.35	0.32	0.28	0.31	1.00

Table II. Tests of Forecast Efficiency

Country	$\hat{\beta}_1$	t-stat (OLS)	t-stat (GMM)
All	0.33	25.03**	11.75**
Austria	0.38	7.22**	5.87**
Belgium	0.31	5.43**	4.68**
Canada	0.30	5.34**	4.04**
Denmark	0.32	5.62**	4.90**
Finland	0.44	7.84**	5.63**
France	0.33	5.99**	4.54**
Germany	0.52	10.55**	7.00**
Ireland	0.21	3.79**	3.15**
Italy	0.41	7.80**	6.49**
Japan	0.25	4.37**	3.69**
Netherlands	0.44	8.37**	5.91**
Norway	0.17	2.92**	2.67**
Portugal	0.29	5.34**	4.79**
Spain	0.42	8.04**	5.74**
Sweden	0.35	6.39**	5.27**
Switzerland	0.38	7.04**	6.02**
UK	0.48	9.44**	6.38**
USA	0.25	4.44**	3.46**

Note: Estimates are based on the regression $r_{i,t,h} = \beta_1 r_{i,t,h+2} + u_{i,t,h}$. GMM t-statistics are computed using the variance covariance matrix given in the text. ** denotes significance at 1 percent level.

Table III. Tests of Independence of Forecast Revision Directions

Country	Frequency / Expected frequency		Chi-Square
	$r_{i,t,h+3} < 0$ and $r_{i,t,h} < 0$	$r_{i,t,h+3} > 0$ and $r_{i,t,h} > 0$	
Canada	92 / 70.8	84 / 54.9	54.9**
France	115 / 84.1	67 / 36.4	108.7**
Germany	110 / 76.1	82 / 47.7	92.9**
Italy	136 / 116.4	42 / 23.4	34.1**
Japan	111 / 83.0	69 / 49.2	43.2**
UK	92 / 62.6	77 / 47.7	98.3**
US	68 / 49.2	110 / 88.2	46.9**

Note: Table presents the number of occurrences satisfying ($r_{i,t,h+3} > 0$ and $r_{i,t,h} > 0$) and ($r_{i,t,h+3} < 0$ and $r_{i,t,h} < 0$) from the 3 × 3 cross table, where the event sets are defined by the signs of forecast revisions: i) positive; ii)zero; and, iii)negative. Expected frequencies are calculated under the hypothesis of independence. Chi-Square statistic is the Pearson Chi-square statistic of independence. The degrees of freedom for the test is (3-1)(3-1)=4. Fisher's Exact Tests were computed but not reported since the results were same. * denotes significance at 5 percent and ** denotes significance at 1 percent level.

Table IV. Steady-state Generalized Variance Decompositions of G-7 Forecast Revisions

Forecast revision :	**Explained by**						
	USA	Japan	Germany	UK	France	Italy	Canada
USA	**91%**	3%	7%	9%	16%	12%	23%
Japan	8%	**87%**	4%	2%	4%	2%	4%
Germany	18%	7%	**84%**	14%	21%	11%	10%
UK	18%	2%	4%	**90%**	3%	3%	10%
France	18%	3%	27%	11%	**76%**	10%	8%
Italy	17%	2%	11%	9%	17%	**89%**	8%
Canada	36%	3%	7%	10%	10%	7%	**80%**

Table V. Speed of Utilization of New Information for G-7 Countries

Months	Canada	France	Germany	Italy	Japan	UK	USA
0	66%	62%	54%	79%	68%	57%	73%
1	81%	72%	65%	84%	73%	70%	87%
2	88%	79%	76%	88%	76%	78%	**94%**
3	**94%**	88%	82%	**93%**	**90%**	86%	98%
4	97%	**92%**	87%	95%	92%	**91%**	99%
5	98%	94%	**91%**	97%	94%	94%	99%
6	99%	95%	94%	98%	97%	96%	100%
7	100%	97%	95%	98%	98%	98%	100%
8	100%	97%	97%	99%	98%	99%	100%
9	100%	98%	98%	99%	99%	99%	100%
10	100%	99%	98%	99%	99%	100%	100%
11	100%	99%	99%	100%	99%	100%	100%
12	100%	99%	99%	100%	99%	100%	100%

Note: Table presents cumulative percentage of the variation in forecast revisions explained by past innovations. Values at month 0 show the immediate use of information.

Figure 1. Generalized Impulse Responses of US Forecast Revisions ±2SE

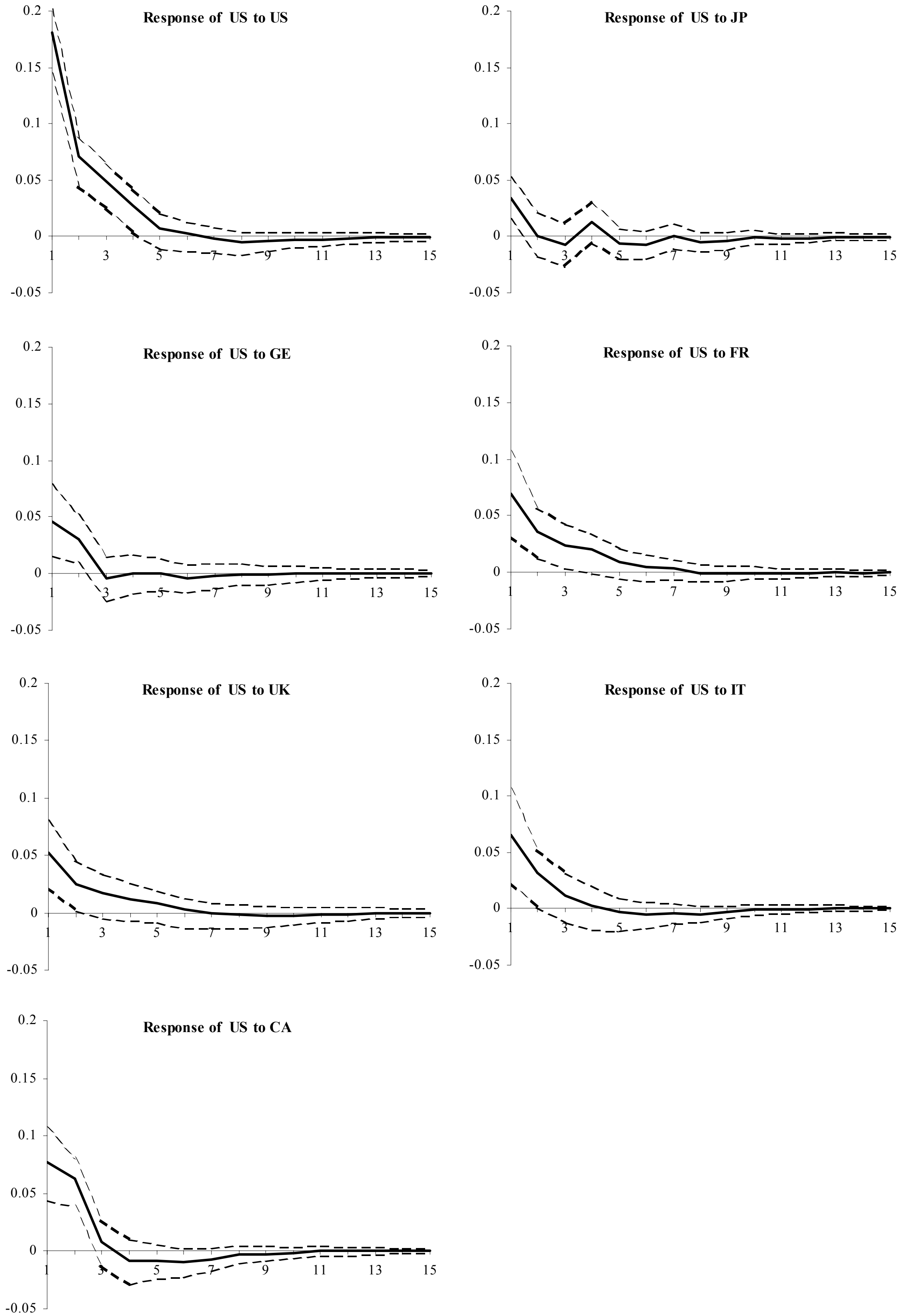

Figure 2. Generalized Impulse Responses of Japanese Forecast Revisions ±2SE

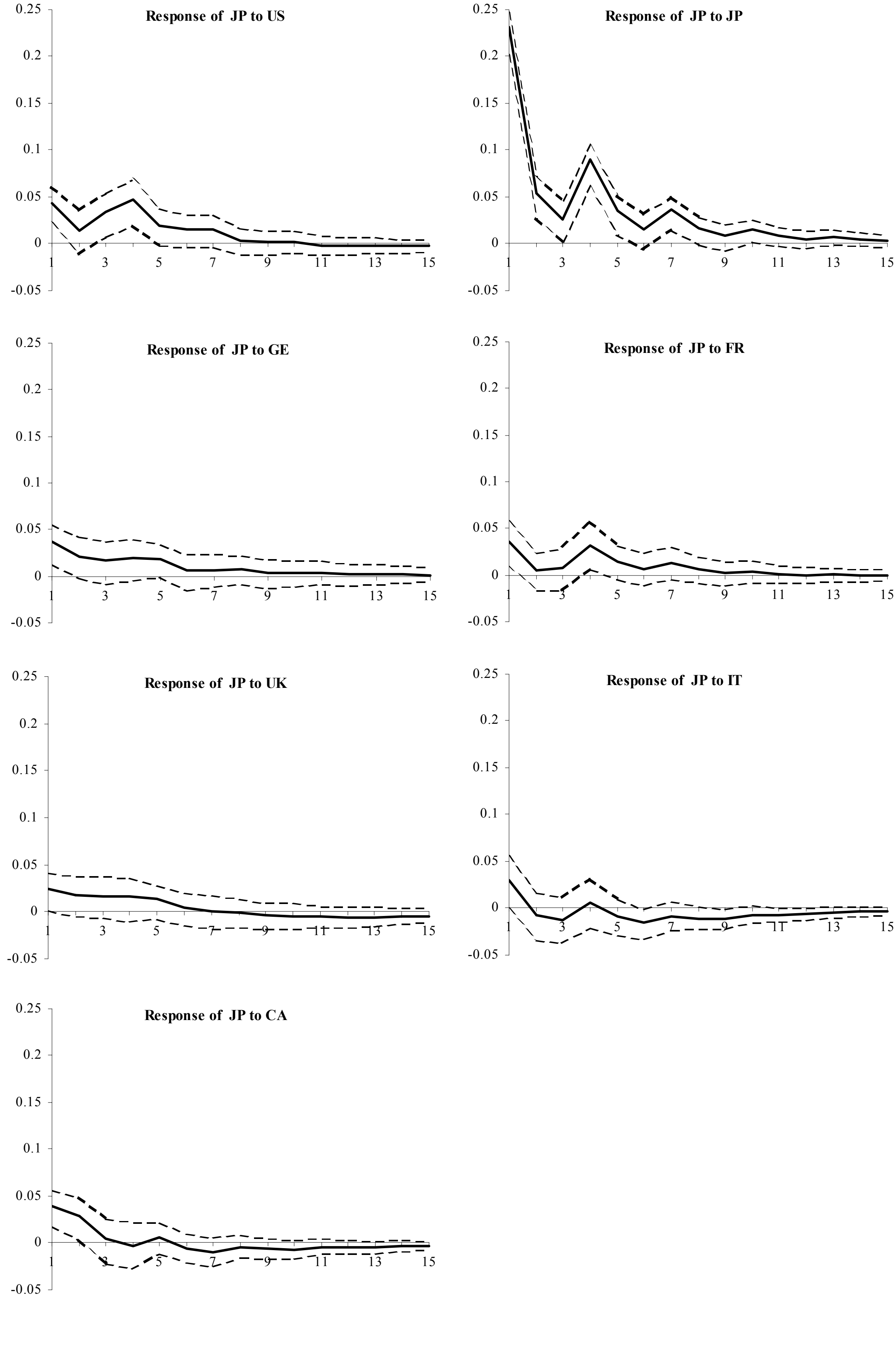

Figure 3. Generalized Impulse Responses of German Forecast Revisions ±2SE

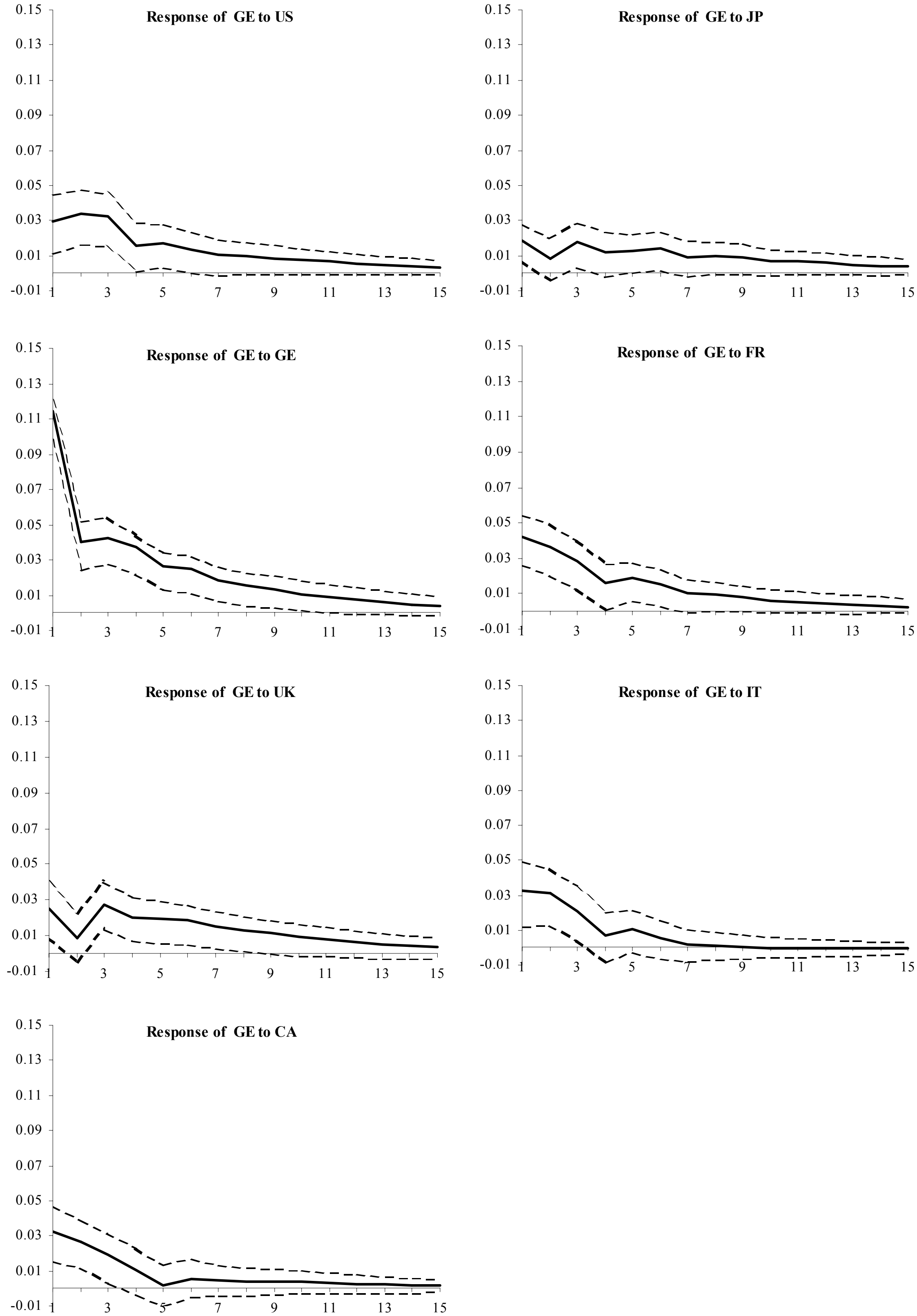

Figure 4. Generalized Impulse Responses of UK Forecast Revisions ±2SE

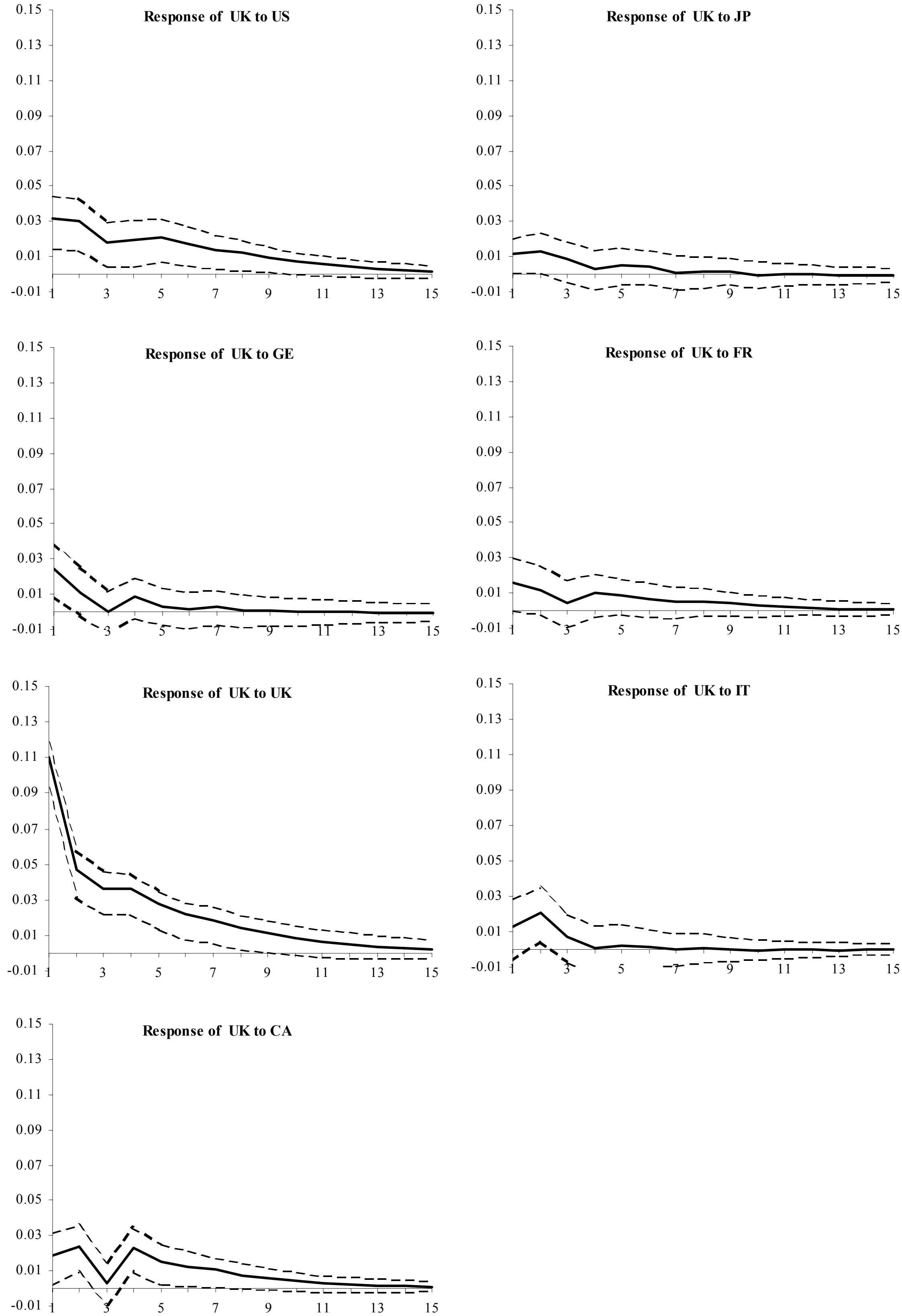

Figure 5. Generalized Impulse Responses of French Forecast Revisions ±2SE

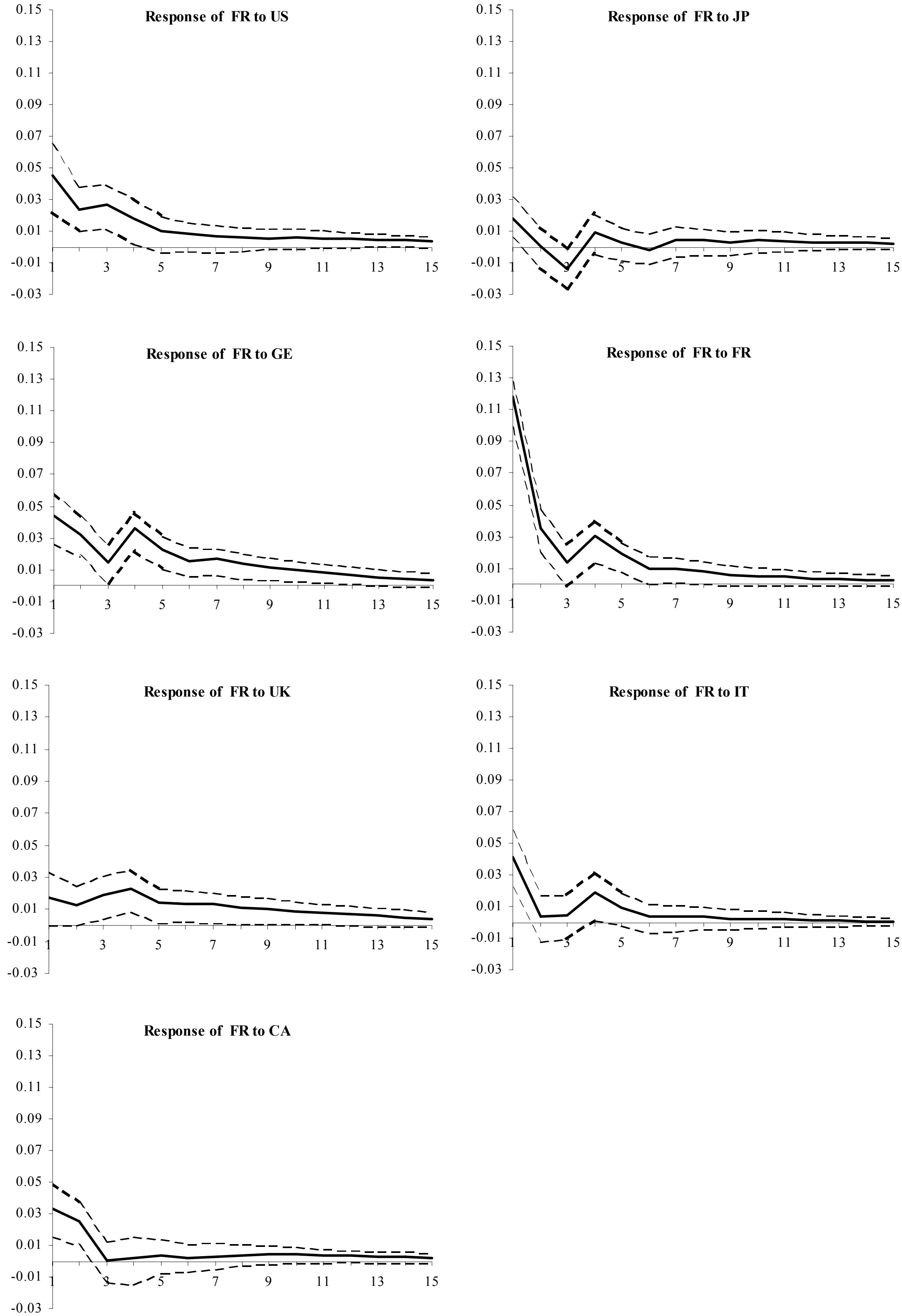

Figure 6 . Generalized Impulse Responses of Italian Forecast Revisions ±2SE

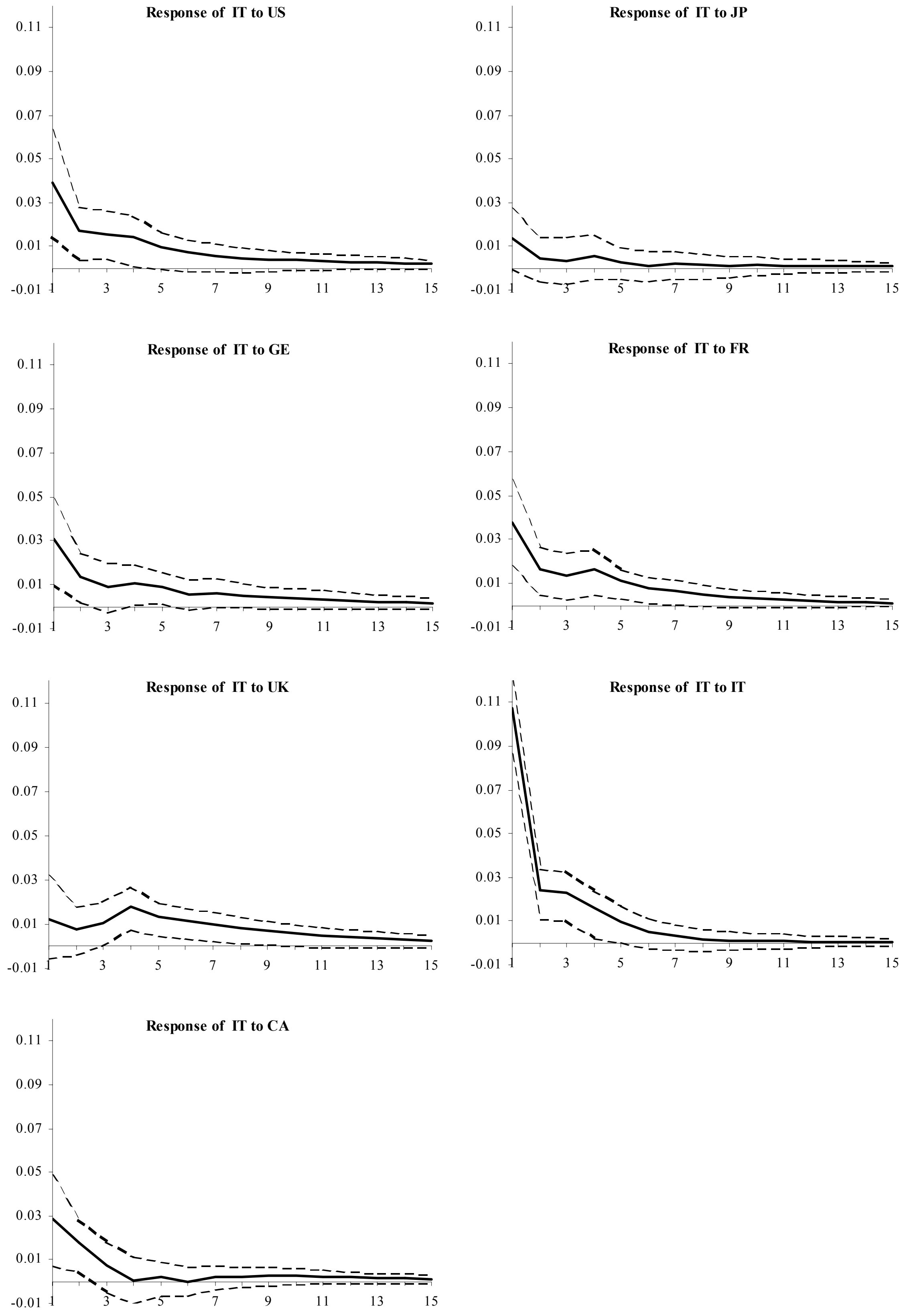

Figure 7. Generalized Impulse Responses of Canadian Forecast Revisions ±2SE