Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 714

December 2001

EXCHANGE RATE FORECASTING: THE ERRORS WE'VE REALLY MADE

Jon Faust, John H. Rogers and Jonathan H. Wright

EXCHANGE RATE FORECASTING: THE ERRORS WE'VE REALLY MADE

Jon Faust, John H. Rogers and Jonathan H. Wright*

Abstract: We examine the forecasting performance of standard macro models of exchange rates in real time, using dozens of different vintages of the OECDs Main Economic Indicators database. We calculate out-of-sample forecasts as they would have been made at the time, and compare them to a random walk alternative. The resulting "time series" of forecast performance indicates that both data revisions and changes in the sample period typically have large effects on exchange rate predictability. We show that the favorable evidence of long-horizon exchange rate predictability for the DM and Yen in Mark (1995) is present in only a narrow two-year window of data vintages around that used by Mark. In addition, approximately one-third of the improved forecasting performance of Mark's monetary model over a random walk is eventually undone by data revisions. Related to this, we find the models consistently perform better using original release data than using fully revised data. Finally, we find that model-based exchange rate forecasts are sometimes better when using Federal Reserve Staff forecasts of future fundamentals instead of actual future values of fundamentals. This contradicts a cherished presumption in the literature that dates all the way back to Meese and Rogoff (1983).

1. Introduction

Economists have never had much luck in forecasting asset prices in general or exchange rates in particular. Even after the fact, analysts are hard-pressed to explain movements in exchange rates through reference to macroeconomic fundamentals. In their classic paper, Meese and Rogoff (1983) demonstrated that, especially at short horizons, a random walk forecast of the exchange rate generally outperforms alternative models drawn from economic theory, including purchasing power parity (PPP), uncovered interest rate parity (UIP), and simple versions of the monetary and portfolio balance models of exchange rates.

Although the Meese-Rogoff finding is remarkably robust, a number of authors have found models whose out-of-sample forecasting performance improves upon a random walk (MacDonald and Taylor (1993), Chinn and Meese (1995), Mark (1995), and MacDonald and Marsh (1997)). While the inference procedures and robustness of results in this wave of post-Meese-Rogoff papers has been called into question (Kilian (1999), Berkowitz and Giorgianni (2001), Berben and van Dijk (1998)), recent work argues that it is still possible—albeit difficult—to beat a random walk forecast with conventional simple specifications. For example, Mark and Sul (2001) use prices, money supplies and output as fundamentals, in a least-squares specification for a panel of countries. Kilian and Taylor (2001) use prices alone.[1]

---

[1]Several recent papers highlight the importance of non-linear adjustment of the exchange rate to the value implied by fundamentals, including Obstfeld and Taylor (1997), Michael, Nobay, and Peel (1997), Taylor, Peel, and Sarno (2001) and Kilian and Taylor (2001). Engel and Hamilton (1990) and Engel (1994) find

Many of the models described above attempt to predict exchange rate movements using economic fundamentals, but there are substantive issues concerning the choice of the appropriate fundamentals data. Meese and Rogoff (1983) base the forecast on the realized values of the fundamentals in the forecast period. By giving the model actual future data in forming the forecast, Meese and Rogoff seem to have given the models an artificial advantage, making it more noteworthy that the models still generally perform worse than a random-walk model. As Meese and Rogoff (1983, p. 12) summarize it, "The structural models in particular fail to improve upon a random walk in spite of the fact their forecasts are based on realized values of the explanatory variables." Of course, any predictability that is found in such a model has limited usefulness to financial market analysts and policymakers who confront the unenviable task of forecasting exchange rates in real time.

Some other authors instead use lagged values of the economic fundamentals to predict future exchange rate movements in an error-correction framework (e.g. Mark (1995)). While this is closer to a real-time experiment than that of Meese-Rogoff, even if such a forecast is evaluated out-of-sample, any measured predictability need not reflect the results that would be obtained by a real-time forecaster. The reason, of course, is that macroeconomic data used in the exercise are often subject to large revisions. Extant work on evaluating exchange rate forecasting models invariably uses the most recent available data, implicitly assuming

that incorporating regime switching – one type of non-linearity – in exchange rate models did not produce forecasting gains over a random walk. Meese and Rose (1991) argue that other forms of non-linearities are largely unimportant for exchange rates. In this paper, we limit explicit consideration to linear models, but note that the issues we raise are relevant to any forecasting model that uses regularly-revised data.

that agents are able to predict future data revisions perfectly. This assumption is not borne out by the data. Faust, Rogers, and Wright (2000) show that revisions to preliminary GDP data are large and are unpredictable for some countries, somewhat predictable for others, but are always far from perfectly predictable.

Although it is common in empirical macroeconomics to work with revised data, there is a growing literature in macroeconomics and finance suggesting that analysis based on real-time data often yields substantially different conclusions from work ignoring data-revisions (Diebold and Rudebusch (1991), Orphanides (2001), Faust, Rogers and Wright (2001), Christoffersen, Ghysels, and Swanson (2001)). Very little work has been done collecting and analyzing real-time data for economies other than the U.S., and hence the importance of real-time data issues for exchange rate forecasting has not, to our knowledge, been examined.

One might suppose that real-time approaches to out-of-sample exchange rate forecasting would only weaken the already limited evidence of predictability in exchange rates. However, Orphanides (2001) and Faust, Rogers and Wright (2001) show that simple monetary policy rules sometimes fit interest rate movements better with real-time data than with revised. Christoffersen, Ghysels, and Swanson (2001) find that real-time macroeconomic data sometimes enters more significantly than revised data in explaining risk pricing in equity markets. In the same way, some exchange rate movements may appear puzzling to us only because we are using data other than those that were actually available to agents at

the time.

In this paper we examine real-time forecasting power of standard exchange rate models based on macroeconomic fundamentals for the Japanese yen, German mark (DM), Canadian dollar, and Swiss franc versus the U.S. dollar during the post-Bretton Woods era. We focus most of our attention on Mark's (1995) monetary model, but the issues we raise are relevant to any forecasting model that uses regularly-revised data, and we show that our main findings also hold in a more general portfolio balance model.

Using an international real-time data set that we constructed, we compare the forecasting performance of the models using different data vintages including using (i) real-time data on the lagged economic fundamentals instead of the *ex post* realized values, and (ii) forecasts of future values of fundamentals instead of the actual future values. Our two most prominent findings are:

1. We conduct the empirical analysis of Mark (1995) for more than 30 data vintages spanning 1988 through 2000. The results reflect what Mark might have found if he had happened to construct his dataset when any particular vintage was the most recent. In most cases, both earlier and later vintages suggest less predictive power than Mark's vintage.[2] This is especially true for the prominent cases of long-horizon predictability of the DM and yen.

2. The predictive power of exchange rate models is nearly uniformly better using real-time fundamentals data than it is using *ex post* revised fundamentals data. This result emerges

---

[2] As discussed below, we deduce that Mark's vintage was from about April 1992

4

consistently using several approaches and models detailed below.

The plan of the remainder of this paper is as follows. In section 2, we describe our international real-time dataset. In section 3, we consider the effects of changing sample period and data revisions on the measured forecasting power of the monetary model and the portfolio balance model. In section 4, we report the result of two real-time forecasting experiments. Section 5 concludes.

2. Description of our International Real-Time Dataset

The relative lack of work in this area can be largely attributed to the difficulties associated with compiling an international real-time data set. We have assembled a quarterly international real-time dataset from historical electronic and paper copies of the OECD's Main Economic Indicators. In the past, the OECD produced an electronic version of MEI up to four times a year: in January, April, July and October. We obtained a number of vintages of the MEI on floppy disks and CDs covering the period from April 1988 to October 2000, inclusive. This gives us a total of 38 vintages of data between 1989 and 2000.[3] The July and October 1994 issues are missing and starting in 1996 only April and October issues were produced. All in all, this leaves us with 38 vintages of MEI, though not all variables are included in all issues.[4]

---

[3]We are grateful to Rita Varley of the OECD's statistics department in Paris for digging through a closet to find these data for us.

[4]The alternative of having a monthly vintage dataset drawn from paper copies of MEI was not viable

For Canada, Germany, Japan, Switzerland and the United States, we extracted real output (GDP or GNP, depending on which was recorded, seasonally adjusted), the CPI price index (not seasonally adjusted), money supply (M1, except M3 for Canada, not seasonally adjusted) and the current account (not seasonally adjusted). These are the variables entering the exchange rate models under study that are regularly revised. We chose these countries following Mark (1995) and used M3 rather than M1 for Canada, for consistency with that paper. Due to problems with the reporting of data in the MEI, for each country some of the 38 vintages are missing.[5] The German results must be interpreted with care due to the effects of German unification.[6] We obtained the exchange rates and short-term interest rates for the four countries from the October 2000 issue of MEI. These data are never revised.

We completed our construction of the real-time data by extracting the Federal Reserve staff forecasts of 4-quarter ahead output for each of these countries from the March, June, September and December Greenbooks, for each year from April 1988 to April 1996, inclusive.[7]

---

because the paper version includes quarterly data only for a few periods prior to the month of issue. The electronic version always contains long quarterly time series as they were observed at the time (generally going back to 1960Q1).

[5] For example, with some regularity, the historical series presented has revised data for the recent period spliced to unrevised data for earlier periods. These appear to simply have been cases where the series was only partially updated by the time the dataset was published.

[6] In vintages through 1994:04, the OECD reports West German GDP. In vintages thereafter, the OECD reports a spliced GDP series which is level-adjusted West German GDP through 1990 and unified GDP thereafter. The level adjustment is meant to produce a meaningful growth rate even in the period of the splice.

[7] The forecasts for individual countries other than the US are not actually in the final version of the Greenbook, but are prepared by the Federal Reserve Staff as part of the process of preparing the Greenbook forecasts. Of course, we have available the staff forecasts right up to the present. Our ending date of 1996 is determined by the Federal Reserve's "five-year window" for disclosing staff projections.

3. Exchange Rate Forecasting with Different Data Vintages

3.1 *Mark's forecasting regression*

Although many approaches have been taken to forecasting exchange rates, we focus on the monetary model as considered by Mark (1995). This is a leading model of exchange rate determination and it uses variables to forecast exchange rates that are subject to substantial revisions. Further, Mark's implementation of the monetary model is one of the most prominent and striking examples of evidence in favor of long-horizon exchange rate predictability.

Let s_t be the log exchange rate, m_t and y_t be log domestic money stock and output and m_t^* and y_t^* be log foreign money stock and output. Following Mark, the money stock variables are constructed as four quarter moving averages, to eliminate seasonality. The fundamental value of the log exchange rate predicted by the monetary model is

$$f_t = (m_t - m_t^*) - (y_t - y_t^*).$$

The basic exchange rate equation at forecast horizon k is,

$$s_{t+k} - s_t = \alpha_k + \beta_k z_t + v_t,$$

where $z_t = f_t - s_t$ and α_k and β_k are coefficients to be estimated. This is an error-correction model.

If the exchange rate is unpredictable then β_k should be equal to zero. However, if the monetary model is useful as a description of the long-run movements in the exchange rate, then β_k should be positive and should initially increase with the horizon k.

7

Mark (1995) estimates this regression on quarterly data for 1973:2 to 1991:4 for the Canadian, Germany, Japanese and Swiss exchange rates, viz-a-viz the U.S. dollar, using horizons k of 1,4,8,12 and 16 quarters. His point estimates of β_k are all positive and tend to increase with the horizon k. While the estimates of β_k are generally not statistically different from zero at, say, the 10 percent level, they are significant in some cases at longer horizons.

Mark also conducts a recursive out-of-sample forecasting exercise using the last 40 quarters of data of his sample for forecast evaluation. Calling the dates of the first and last available observations 1 and T respectively, the method for each k and each currency is as follows. Estimate the parameters α_k and β_k using observations 1 through $T - 40$, and construct a forecast for the k-period exchange rate change starting in $T - 40$. Next extend the end of the estimation sample to $T - 39$ and forecast the k-period change from $T - 39$, and so on, until the estimation sample ends in $T - k$. This results in 40 $k = 1$ quarter forecasts and 25 $k = 16$ quarter forecasts.

Mark reports the ratio of the root mean square error (RMSE) of these forecasts to that of a random walk forecast (without drift).[8] Ratios less than one mean that the monetary model is outperforming the random walk. Mark also reports bootstrap p-values for a one-tailed test of the null hypothesis that this ratio is one versus less than one.

---

[8]It is a familiar fact in this literature that the random walk without drift outperforms the random walk with drift in forecasting these exchange rates over this period: the random walk without drift is a harder model to beat.

Mark reports several striking conclusions. While the monetary model is generally not much better than the random walk at short horizons, the predictive power seems to rise with forecast horizon until at the 16-quarter horizon the relative RMSE for the monetary model for Germany, Japan, and Switzerland are 0.52, 0.57, and 0.41, respectively. Canada is the exception to the rule with a relative RMSE of 1.70 at the 16 quarter horizon.

Although the point estimates of the relative RMSE at long-horizons are very impressive, the statistical significance of the results is not so clear. This is as one should expect given the fact that the data are persistent and that one only has 2 1/2 nonoverlapping 16-quarter forecasts in a 40-quarter forecast period. In any case, Mark rejects that the relative RMSE is one at the 10 percent level or better for horizon 12 for the DM and for horizon 16 for the DM, and Swiss franc.[9] Mark also emphasizes that while it is difficult to reject the hypothesis that the monetary model is no better than a random walk, the statistical significance tends to rise with horizon.[10] These results are important evidence against the view that the link between fundamentals and exchange rates is inscrutable. For later purposes it is important to note that we are able to reproduce Mark's results to a very high degree of accuracy using our April 1992 vintage.[11]

3.2 *The monetary model using other data vintages*

We first repeat Mark's exercise using each of our vintages of data.[12] We repeat the recur-

---

[9] We refer here to column iv, Table 4 of Mark (1995).

[10] That is, p-values for rejecting the null hypothesis of 1 for the relative RMSE fall with horizon.

[11] This is the first vintage that has all the data the Mark used.

[12] For this exercise, we start with the July 1989 vintage in order to have enough data to estimate β_{16} while

sive exercise with each vintage, in each case keeping the last 40 quarters for out-of-sample forecasting. Thus the 40-quarter forecast period moves forward in time as we move to later vintages. Further, the estimation period, which begins in 1973:2 in every case, grows as we move to later vintages. In the end, where Mark has one relative RMSE for a each forecast horizon and currency, we have a time series of relative RMSEs dated by when the relevant data were released.

While we report the Canada and Switzerland results throughout, we focus on the results for Germany and Japan. As in Mark, the monetary model performs badly for Canada. The issues we discuss are second order relative to this gross failing. For Switzerland, the monetary model performs well in Mark: the relative RMSE is below 1 for all horizons, and at horizon 16, one rejects that the relative RMSE is one at the 2 percent level. Except for the earliest vintages, the effects of changing sample period and vintage do not greatly alter this picture for Switzerland.

The relative RMSEs for each currency and horizon are shown in Figure 1. For Germany and Japan, there is a distinct U-shape to the relative RMSEs, for $k=4$, 8, 12 and 16. The relative RMSEs first fall, reaching a minimum sometime around the April 1992 vintage, and then rise, sharply for Germany.

Figure 2 gives the p-values associated with the relative RMSEs in Figure 1. In particular, these are p-values for one-tailed tests of null hypothesis that the relative RMSEs in

_____

holding back 40 quarters of data for forecasting. Given the missing vintages for various countries, this means we end up with 16, 24, 32, and 13 vintages for Canada, Germany, Japan, and Switzerland, respectively.

Figure 1 are equal to 1 versus being less than one. Here and throughout this paper, p-values are obtained by a bootstrap as described by Mark.[13] We also computed p-values using the bootstrap methodology proposed by Kilian (1999), but, to conserve space, we do not report these.[14]

For Germany and Japan, the p-values in Figure 2 show an even more pronounced trough around the time of the April 1992 vintage than did the relative RMSEs. Looking at horizon 16, we see that the Japan and Germany p-values are both less than 0.15 only in vintages between 91:04 through 93:01, inclusive. For the Switzerland p-value to also be less than 0.15 the vintage range would have to be 91:07 through 93:01. Thus, there was less than a 2-year window in which the data could be drawn if one wanted to find long-horizon p-values approaching usual significance levels.

The forecasting results can be understood a bit better by looking at the estimates of β_k using the data from 1973:2 onward in each vintage (Figure 3). The time series of p-values associated with the one-tailed t-test of $\beta_k = 0$ are given in Figure 4.[15]

The estimates of β_k and the associated p-values vary considerably across data vintages.

---

[13]We use what Mark (1995) refers to as the nonparametric bootstrap. The p-values are one-sided, testing the hypothesis that relative RMSE is one against the alternative that it is less than one. In all implementations of the bootstrap, we use 2,000 bootstrap replications.

[14]These p-values, available from the authors on request, are generally lower than those using the bootstrap of Mark (1995), but exhibit the same patterns. Because the Kilian p-values are more favorable to the monetary model than the Mark values, it is worth noting that even by the Kilian p-values one cannot reject that the relative RMSEs are 1 at horizon 16 in recent vintages at the 10 percent level.

[15]The overlapping structure of the data leads to substantial serial correlation in the error term for $k > 1$. Accordingly, throughout the paper, all t-statistics are autocorrelation robust using a Bartlett window with a truncation lag of 20, following Mark (1995) (results with a data dependent truncation lag are similar and available from the authors on request).

For Germany and Japan, the estimates of β_k for horizons greater than 1 show a clear pattern of falling in later data vintages, which is consistent with the weakening predictive power of the monetary fundamentals shown in Figures 1 and 2. In the case of Germany, the associated p-values get larger in later data vintages, moving clearly outside the range of being statistically significant. In the case of Japan, the associated p-values first rise and then fall, reflecting the offsetting effects of a falling point estimate and an increasing estimation sample size.

3.3 *The pure data revision effect*

The finding that most data vintages give less evidence of predictability of exchange rates than is found with the April 1992 vintage mixes two distinct effects. The first effect is due to the fact that the forecast period—the final 40 observations—changes with vintage. The second effect arises because for a fixed sample period, the data in each vintage may differ due to revisions—the pure data revision effect.

The finding that exchange rate forecastability is sensitive to the sample period is not a new result. Although Mark speculated that a longer sample period would strengthen the finding of predictability in exchange rates, Kilian (1999) and Groen (1999) both showed that adding the 1990s data weakens Mark's result. We are not aware of work documenting the pure revisions effect on the predictability of exchange rates.

Data revisions can have large effects on the fundamentals. For example, Figure 5 plots the revisions to the log money stock and log output between the April 1992 vintage and the most recent available vintage for Canada, Germany, Japan, Switzerland and the United

States.[16] These revisions can affect the low frequency properties of these time series. Notice in particular that the growth rate of U.S. output has been revised upwards over time, while the growth rate of Japanese output has been revised downwards over time. The pessimistic bias in initial releases of U.S. output growth data and the optimistic bias in the case of Japan were noted in Faust, Rogers and Wright (2000).

In order to illuminate the pure data revisions effect on exchange rate predictability, we once again repeat Mark's exercise using our vintage data, but this time we hold the estimation and forecast samples to be the same as in the Mark paper. We perform this exercise on all vintages from April 1992 onward (earlier vintages do not have all the data to perform the exercise using Mark's sample dates).

Figures 6–9 mirror figures 1–4, giving relative RMSEs, associated p-values for a test that the relative RMSE is 1, estimated β_ks, and associated $p-$values for a test that $\beta_k = 0$, respectively. While magnitudes vary, the direction of effects is extremely similar across countries and horizons. With few exceptions, relative RMSEs rise as the data are more revised, as do the associated p-values; estimated β_ks fall as the data are revised, and the p-values rise. Once again it should be noted that effects for Canada are of second order importance relative to the gross failure of the model; the revisions effect for Switzerland fit the pattern, but are small in magnitude.

For Germany and Japan, however, the pure revisions effect is substantial. The April

---

[16]For Canada, the most recent available vintage is the October 1996 vintage because MEI subsequently ceased to report M3 data. For all other countries it is the October 2000 vintage.

1992 vintage gives the best relative RMSEs (Figure 6), and these rise steadily as the data are revised. For the longest horizon, using a more recent data vintage increases the relative RMSE from 0.52 to 0.64 for Germany and from 0.55 to 0.69 for Japan. While the estimated relative RMSEs remain less than one, the p-values for testing the null hypothesis that the relative RMSEs are actually one, rise above the 10 percent level in each case. Thus, revisions alone have a modest but important effect on the estimated relative RMSE, and leave the Germany and Japan long-horizon forecasting power statistically insignificant at conventional levels.

3.4 *Portfolio Balance Model*

To see if our results for the monetary model hold in models that include other regularly revised macroeconomic variables, we now repeat a subset of the exercises above using the portfolio balance model. As noted by Frankel (1982), Hooper and Morton (1982), and Meese and Rogoff (1983), a prototype version of this model posits that the nominal exchange rate is determined by home-foreign differentials in the monetary fundamentals used above as well as short-term interest rates, expected inflation rates, and cumulated current account balances.

We supplement our data set with overnight interest rates, actual CPI inflation, and the balance on current account in local currency. All variables are taken from various vintages of the MEI. Of these series, only the current account is subject to regular revisions, which for all of our countries are large and frequent. For estimation and forecasting, we retain the long-horizon error correction specification, equation (2) above, but now z_t is the deviation

14

of the spot exchange rate from the "portfolio balance fundamentals." This approach follows

Chinn and Meese (1995).

Consider the model for the yen-dollar exchange rate. In this case, the portfolio balance

fundamental is $(m - m^*) - (y - y^*) + 3.0(\pi - \pi^*) + 0.5(i - i^*) - 0.0001(b - b^*)$, where i

is the short rate, π is the 4-quarter ahead change in the CPI, and b is the accumulated

current account balance in local currency. The parameters in this expression are quite close

to those used by Chinn and Meese (1995) and to those obtained by estimating a dynamic

OLS regression (Stock and Watson (1993)).

To begin, we use the April 1992 vintage data and the same out-of-sample forecasting

exercise used to replicate Mark's results for the monetary model.[17] At the 16-quarter horizon

for Japan, we find that the ratio of RMSE of the portfolio balance model to random walk

is 0.39, which is about 25 percent better than the performance of the monetary model.

Using the same construction of the portfolio balance fundamental, same error-correction

specification, but now using the most recent available vintage dataset and most recent 40-

quarter forecast period, we find that the relative RMSE is 1.71. As noted above, this decline

in predictive capability is due to both changing the estimation and forecast sample periods

and due to a pure data revisions effect.

To parse out the pure data revisions effect, we use the most recent vintage but the

same 40-quarter forecast period used by Mark. In this case, we find that the relative RMSE

---

[17]In the vintage exercises below, we hold fixed these parameter estimates defining the fundamentals. We continue to estimate α_k and β_k in a rolling manner just as in the monetary model exercises.

is 0.63. The degradation of forecasting power due to the pure revisions effect is even more substantial than observed for the monetary model. One might have wondered if the results for the monetary model were due to some peculiarities of how GDP has happened to be revised between the late 1980s and the present. This result suggests that the phenomenon may be more general.

4. Real-time Out of Sample Forecasting

None of the exercises in section 3 directly measure real-time forecasting performance of the models in question. A working forecaster uses all data that have been released up to the current time, say, T in forming a forecast for $T + k$. In our forecasting experiments, we have up to now followed Mark in reserving the final 40 observations for forecasting. In this section we report results for two real-time forecasting exercises, one using the monetary model and one using a real-time variation on the classic Meese-Rogoff approach.

4.1 *A real-time test of the monetary model*

In this experiment, we retain the basic framework from section 3, but evaluate the real-time forecasts of the model using all available data at each point in time. That is, for each vintage, we use all the available data to estimate the α_k and β_k, and then we form a single k-period forecast using those estimates and the final observation for the fundamentals data in the vintage.

For purposes of comparison, we also consider a revised data, out-of-sample forecasting

experiment in which we keep the estimation and forecast periods just as in the experiment just described, but we substitute the most recent available vintage of data in all cases. Most out-of-sample forecasting experiments in the existing literature use the most recent available vintage of data, and this will provide a basis of comparison.

The relative RMSEs of the monetary model forecast are shown in Table 1 for all countries, forecast horizons, and for both the real-time forecasting and revised data forecasting experiments. For both experiments, the relative RMSEs are generally greater than 1, and increase with the horizon, for Germany, Japan and Switzerland. This reflects the breakdown of the monetary model over this sample period.

It is notable, however, that for all countries and horizons (except Canada at horizons 12 and 16), the relative RMSE is lower in the real-time experiment than it is in the revised data experiment. These differences are statistically significant in many cases. While real-time data does not rescue the monetary model over this sample period, it is striking that the out-of-sample predictive power of the monetary model is nearly uniformly better using real-time fundamentals data than it is using *ex post* revised fundamentals data.

4.2 *Meese-Rogoff in Real-Time*

Meese and Rogoff (1983), in work that predates the cointegration literature, forecast exchange rates by simply regressing the exchange rate on the macroeconomic fundamentals and then using these parameter estimates and the *ex post* realized and revised values of the future economic fundamentals to predict the future exchange rate. These fundamentals data

would not have been available to agents forecasting in real time. In this section we evaluate the magnitude of this artificial advantage.

For each vintage from April 1988 to April 1996 inclusive we regress s_t on f_t using all data in that vintage from 1973:2 onwards. We consider a Meese-Rogoff experiment and a real-time variant of that experiment. In the Meese-Rogoff experiment, we construct the 4-quarter ahead forecasts using those parameter estimates and the observations on the *actual* 4-quarter ahead fundamentals obtained from the most recent available data vintage. In the real-time variant, we construct 4-quarter ahead forecasts, using those parameter estimates and the 4-quarter ahead *forecast* of the fundamentals, obtained from the Federal Reserve Greenbook forecasts.[18]

In the Meese-Rogoff experiment, the relative RMSEs are 2.17, 1.81, 1.78 and 2.21, for Canada, Germany, Japan and Switzerland, respectively. The model did substantially less well than a random walk. In the real-time variant, the relative RMSEs are 0.98, 2.15, 1.60 and 3.21, for Canada, Germany, Japan and Switzerland, respectively. Thus using the real-time forecasts instead of the unrealized future data can dramatically affect the forecasting performance, but the direction of the effect is ambiguous. The real-time forecasts are better for Canada and Japan, and worse for Germany and Switzerland. Indeed while the monetary model generally predicts much less well than the random walk in these experiments, for

---

[18]In this subsection we consider only 4-quarter ahead forecasts because the fundamentals forecasts are not available at longer horizons. We only have forecasts of output, not money stock. For money stock we use the ex-post realized and revised future data, implicitly assuming that agents can perfectly predict the money supply 4 quarters ahead.

Canada, the monetary model just beats a random walk in the real-time experiment (though this is not statistically significant). The conventional wisdom is that Meese and Rogoff were giving the economic model an artificial advantage by giving the forecaster the *ex post* realized and revised future fundamentals data. This need not necessarily be the case.

5. Conclusion

We perform several variations of experiments all intended to assess the impact of data revisions and sample period on exchange-rate forecastability. We find that measured forecastability is quite sensitive to these issues.

Had Mark (1995) constructed his dataset at almost any time other than early 1992, he would have found much less evidence of predictability. There is only a 2-year window around Mark's vintage in which the monetary model beats the random walk for both the DM and yen at even the modest 15 percent level. That results like this arise from time to time is perhaps not surprising in an area where many researchers independently fit many models to each successive vintage. Only significant results are published. This sort of process seems likely to raise familiar data-mining problems at the profession-wide level, even though any individual researcher at any point in time will always simply use use the most recent vintage of data.

The fact that most other vintages give much less evidence of exchange-rate forecastability is a consequence of both changing sample period and a pure revision effect. For a fixed

19

time period, the monetary and portfolio balance models we examine consistently perform better using original release data than using revised data. We reach this conclusion in a number of different experiments. This is consistent with some recent work of Christoffersen, Ghysels, and Swanson (2001) regarding using macoreconomic factors to model equity returns. Giving the exchange-rate forecasting model the supposed benefit of knowing the final revised data turns out to be more of a hindrance than a help. If data revisions are not forecastable and if agents price assets based on the data available to them in real time, then it would be expected that using revised data for predictors would reduce forecastability.

References

Berben, R.B. and D.J.van Dijk (1998): Does the absence of cointegration explain the typical findings in long horizon regressions?, Econometrics Institute, Erasmus University Rotterdam, Report 9814.

Berkowitz, J. and L.Giorgianni (2001): Long-horizon exchange rate predictability?, *Review of Economics and Statistics*, 83, pp.81–91

Chinn, M.D. and R.A.Meese (1995): Banking on currency forecasts: how predictable is change in money?, *Journal of International Economics*, 38, pp.161–178.

Christoffersen, P., E. Ghysels and N.R.Swanson (2001): Let's Get "Real" about Using Economic Data, mimeo.

Diebold, F.X. and R.Mariano (1995): Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, pp.253–262.

Diebold, F.X. and G.Rudebusch (1991): Forecasting output with the composite leading index: a real-time analysis, *Journal of the American Statistical Association*, 86, pp.603–610.

Engel, C. and J.Hamilton (1990): Long swings in the dollar: are they in the data and do markets know it? *American Economic Review*, 80, pp.689–713.

Engel, C. (1994): Can the markov-switching model forecast exchange rates? *Journal of International Economics*, 36, pp.151-165.

Faust, J., J.H. Rogers and J.H.Wright (2000): News and noise in G-7 GDP announcements, International Finance Discussion Paper 680, Federal Reserve Board, Washington, D.C.

Faust, J., J.H. Rogers and J.H.Wright (2001): Forward-looking monetary policy rules in an international real-time data set, mimeo, International Finance Division, Federal Reserve Board, Washington, D.C.

Frankel, J.A. (1982): The Mystery of the Multiplying Marks: A Modification of the Monetary Model, *Review of Economics and Statistics*, 64, pp.515-519.

Groen, J.J.J. (1999): Long Horizon Predictability of Exchange Rates: Is It for Real? *Empirical Economics*, 24, pp.451–469.

Hooper, P. and J.E. Morton (1982): Fluctuations in the Dollar: A Model of Nominal and Real Exchange Rate Determination, *Journal of International Money and Finance*, 1, pp.39-56.

Kilian, L. (1999): Exchange rates and monetary fundamentals: what do we learn from

long-horizon regressions? *Journal of Applied Econometrics*, 14, pp.491–510.

Kilian, L. and M.P.Taylor (2001): Why is it so difficult to beat the random walk forecast of exchange rates?, mimeo, University of Michigan and University of Warwick.

MacDonald, R. and M.P.Taylor (1993): The monetary approach to the exchange rate: rational expectations, long-run equilibrium, and forecasting, IMF Staff Papers 40.

MacDonald, R. and I.W. Marsh (1997): On Fundamentals and Exchange Rates: A Casselian Perspective, Review of Economics and Statistics, 79, pp. 655–64.

Mark, N.C. (1995): Exchange Rates and Fundamentals: Evidence on Long-Horizon Prediction, *American Economic Review*, 85, pp.201–218.

Mark, N.C. and D.Sul (2001): Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel, *Journal of International Economics*, 53, pp.29–52.

Meese, R. and K.Rogoff. (1983): Empirical exchange rate models of the seventies: do they fit out of sample?, *Journal of International Economics*, 14, pp.3–24.

Meese, R. and A.K.Rose (1991): An empirical assessment of non-linearities in models of exchange rate determination, *Review of Economic Studies*, 58, 603–619.

Michael, P., A.R.Nobay and D.A.Peel (1997): Transactions costs and non-linear adjustment in real exchange rates: an empirical investigation, *Journal of Political Economy*, 105, pp.862–879.

Obstfeld, M. and A.Taylor (1997): Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited, *Journal of the Japanese and International*

Economies, 11, pp.441–479.

Orphanides, A. (2001): Monetary policy rules based on real-time data, *American Economic Review*, forthcoming.

Stock, J. and M.Watson (1993): A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica*, 61, pp.783-820.

Talyor, M., D.Peel and L.Sarno (2001): Non-linear mean-reversion in real exchange rates: towards a solution to the purchasing power parity puzzles, *International Economic Review*, forthcoming.

Table 1: Relative RMSE of monetary model in real-time and revised data out-of-sample forecasting

k	data	Canada	Germany	Japan	Switzerland
1	rt	0.99*	1.01	1.00	1.00
	rv	1.02*	1.01	1.01	1.02
4	rt	0.81	1.10	0.98*	0.93
	rv	0.92	1.11	1.02*	0.93
8	rt	0.63	1.87*	1.01*	1.02
	rv	0.69	1.99*	1.08*	1.09
12	rt	0.65*	2.52*	1.17*	1.01
	rv	0.58*	2.77*	1.26*	1.12
16	rt	0.55*	2.90*	1.40*	1.13
	rv	0.50*	3.26*	1.53*	1.30

Notes: The entries in this table show the ratio of the out-of-sample RMSE from the monetary model to that from the random walk model in the real-time (data 'rt') and revised (data 'rv') data experiments described in the text. For each country, the two models are used to forecast exchange rates at horizons of $k=1,4,8,12$ and 16. The entries marked with an asterisk refer to cases in which the difference in RMSE between the real-time and revised-data experiments is statistically significant at the 5% level, using the test statistic proposed by Diebold and Mariano (1995) (autocorrelation robust, with a lag length parameter of 20).

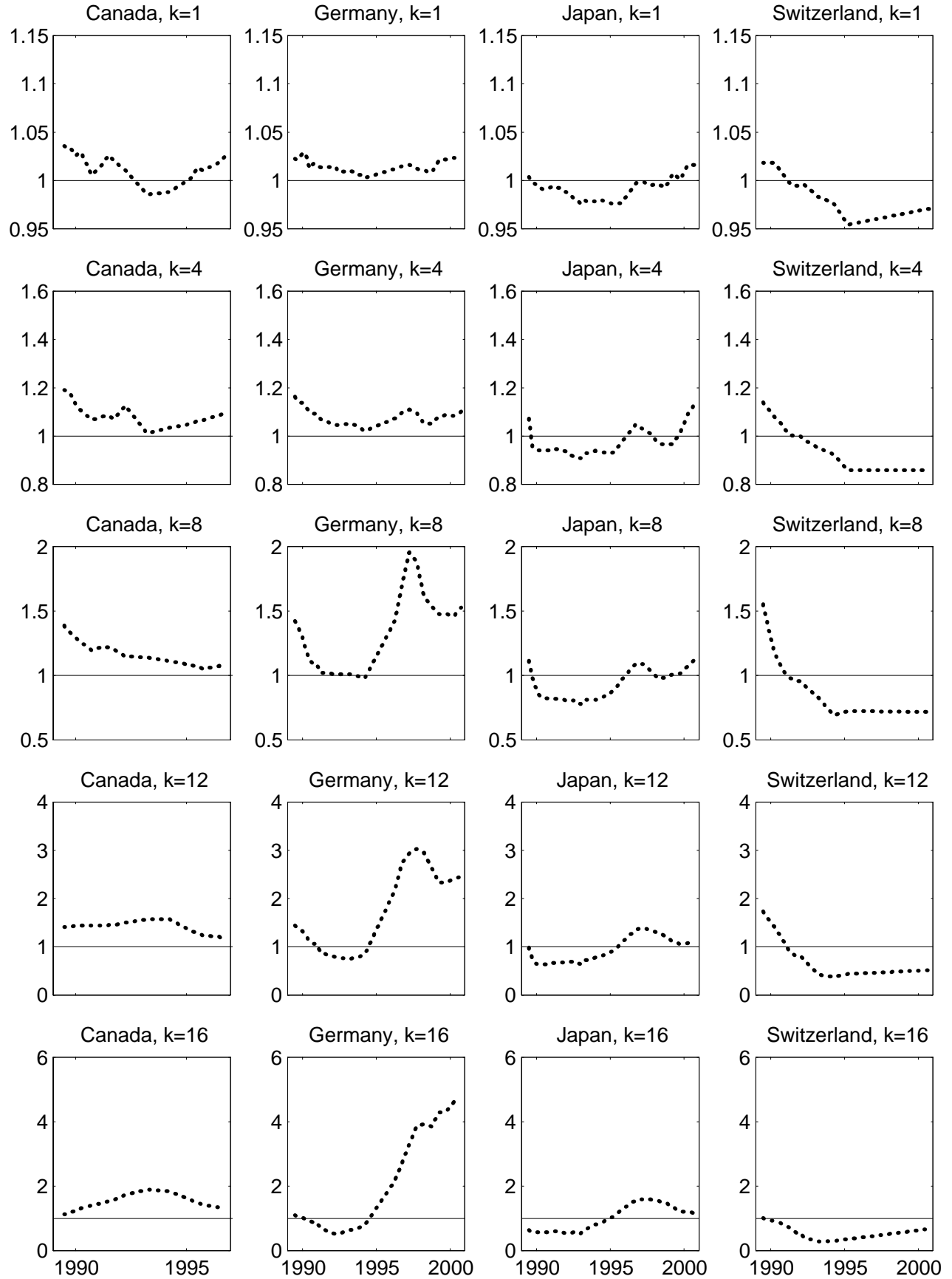Figure 1: Out-of-sample relative RMSE with different data vintages

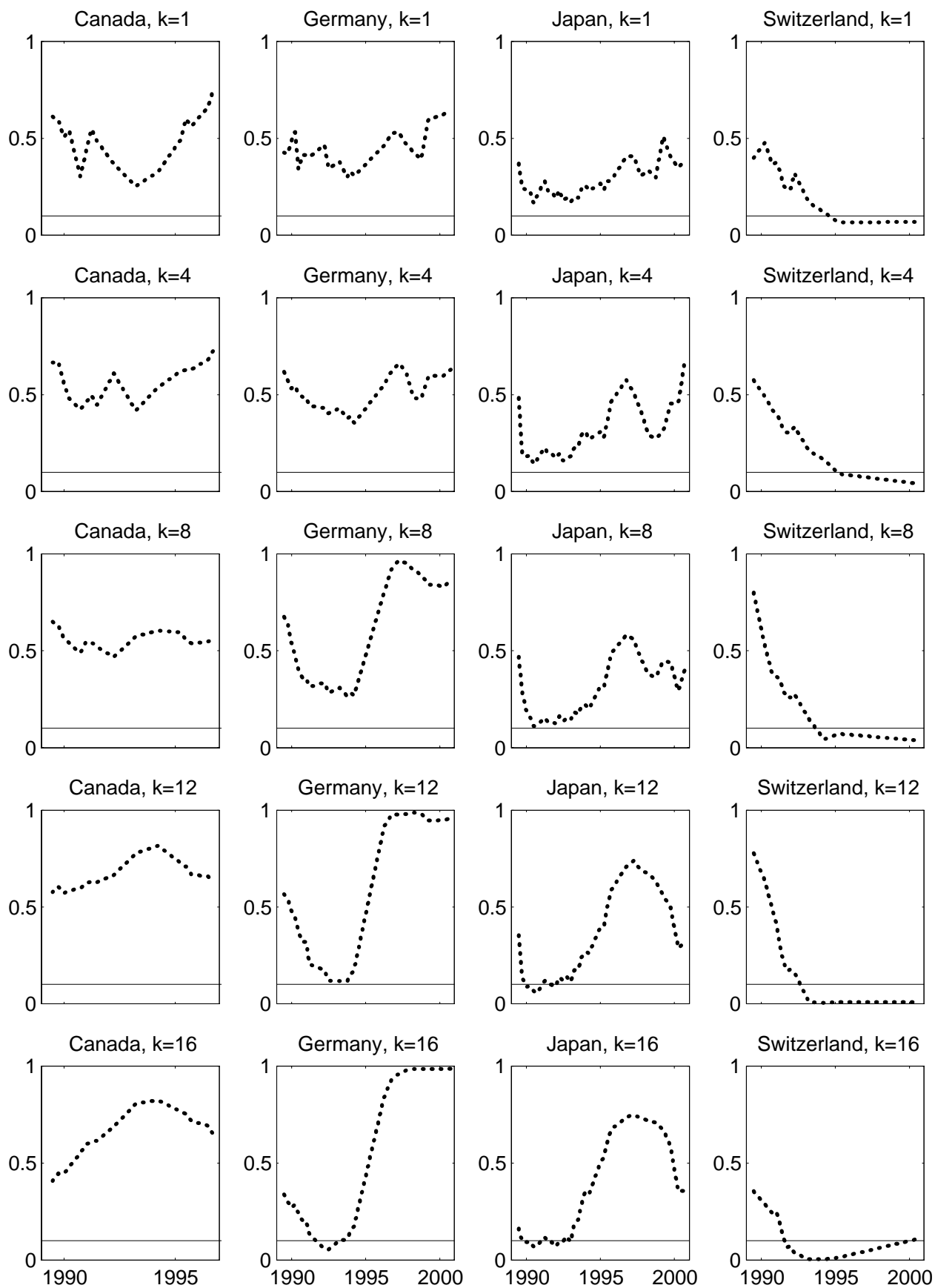Figure 2: Bootstrap p-values for out-of-sample relative RMSE using different data vintages

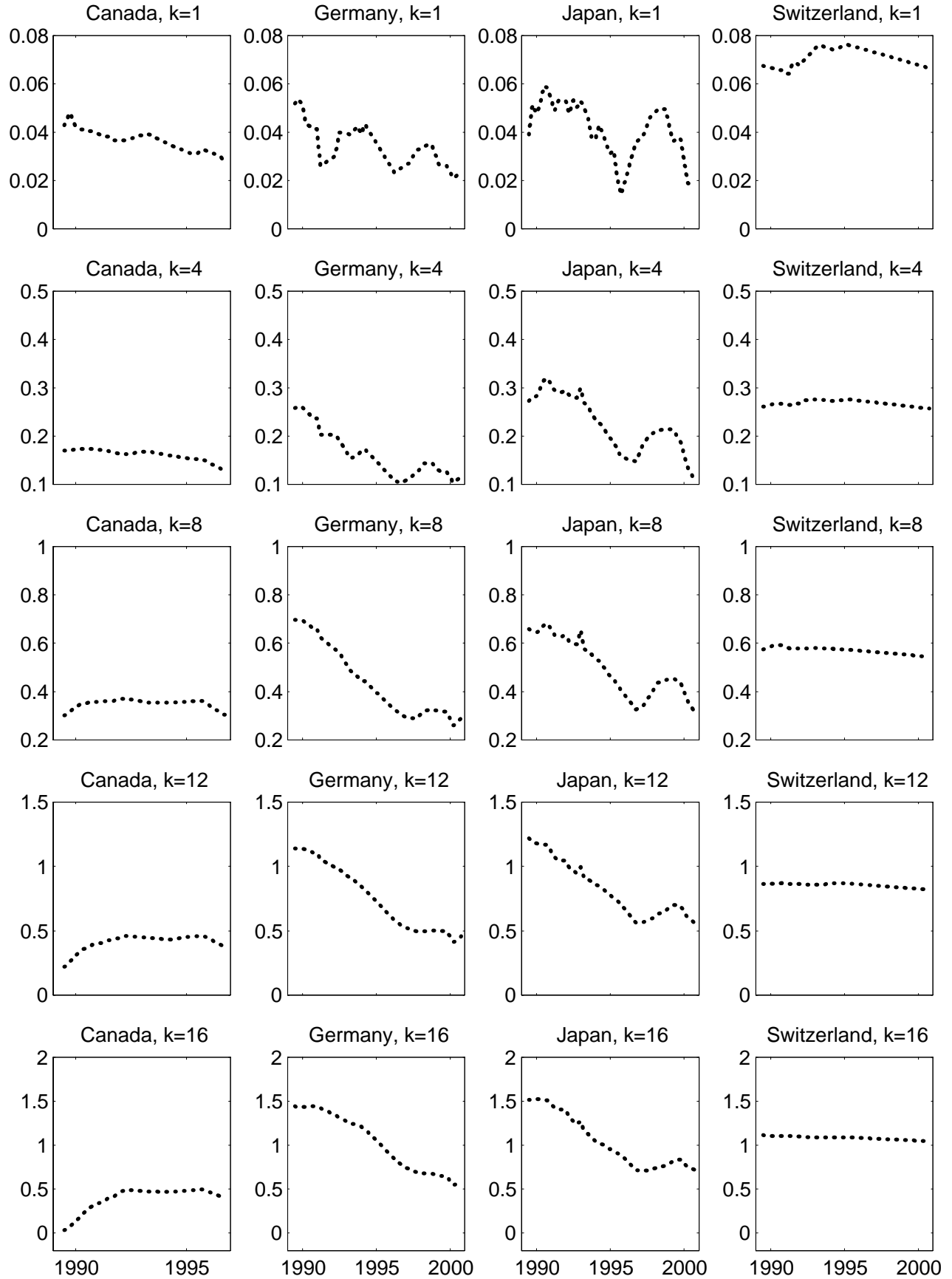Figure 3: Estimates of β_k using different data vintages

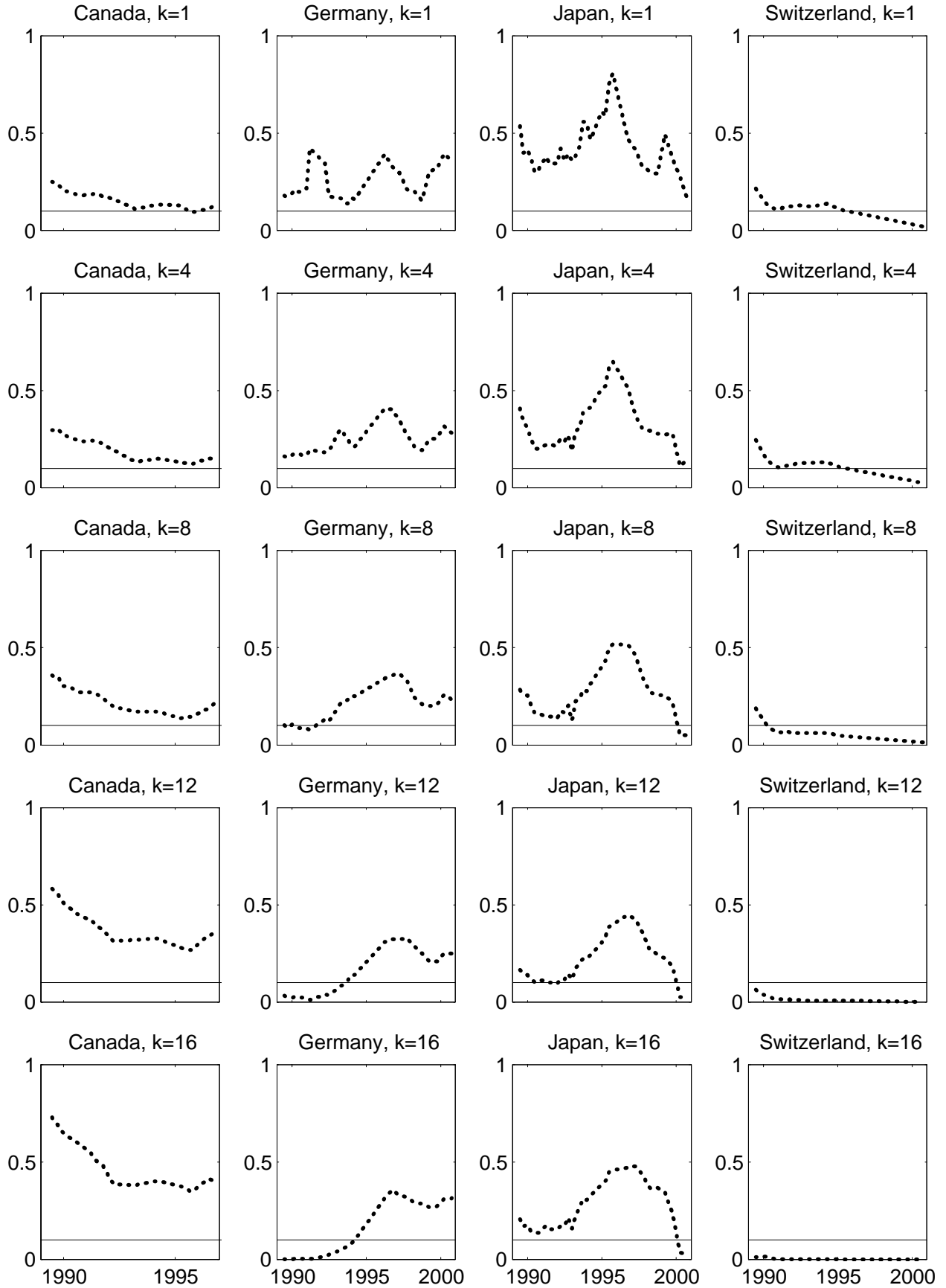Figure 4: Bootstrap p-values for t_k using different data vintages

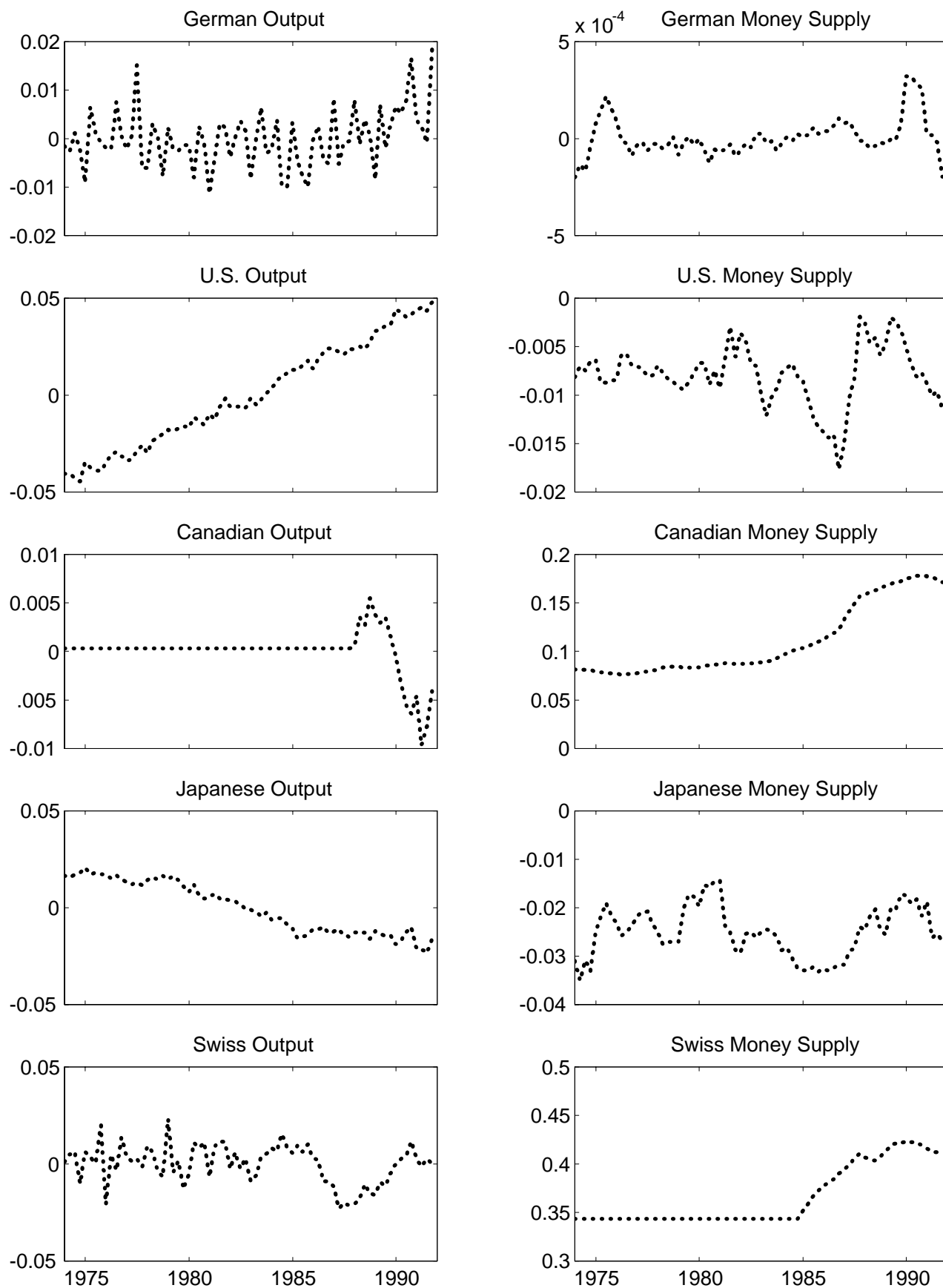Figure 5: Revisions to log output and money stock from April 1992 to latest release

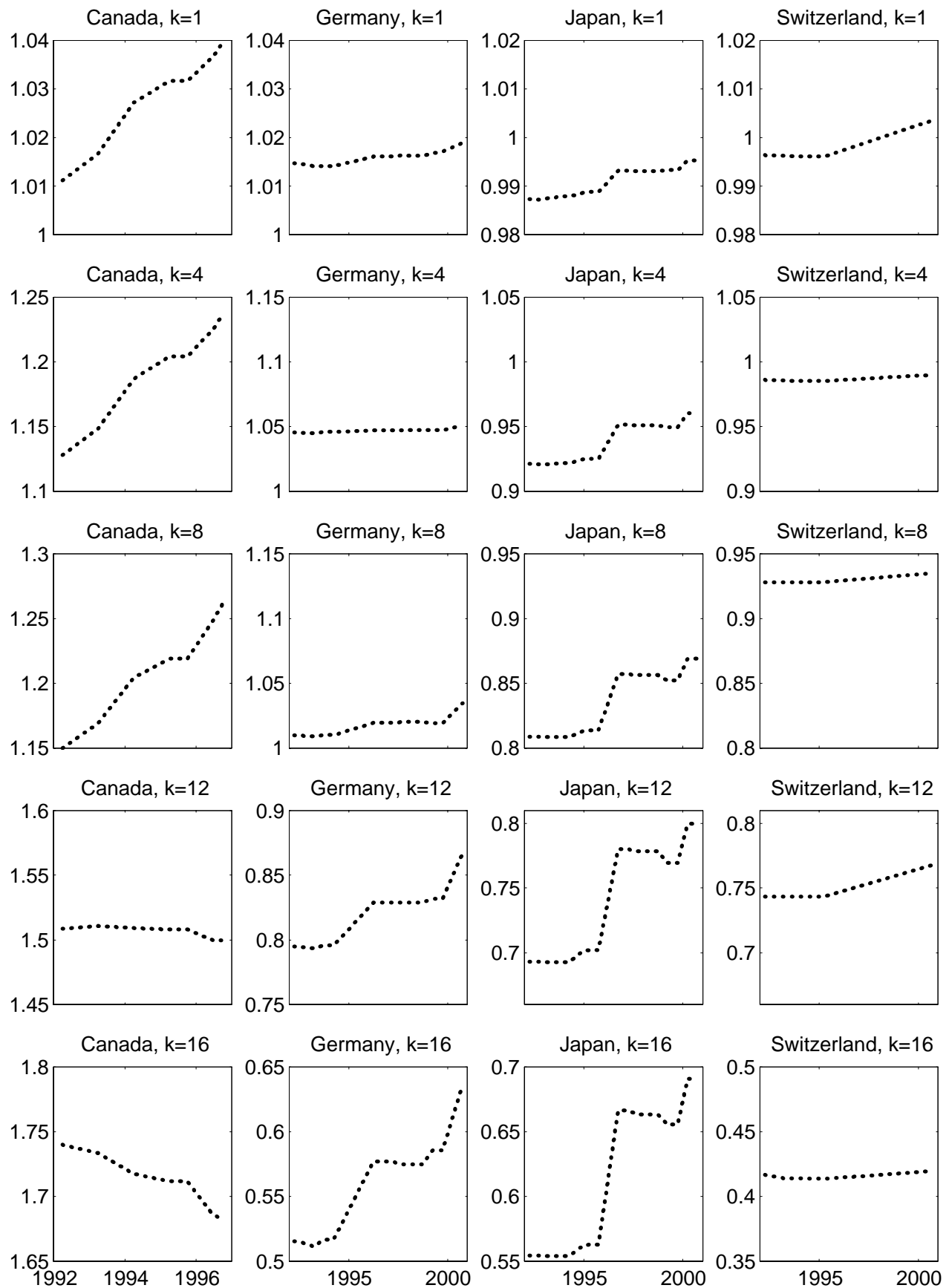Figure 6: Out-of-sample relative RMSE using different data vintages (Mark's sample period)

Figure 7: Bootstrap p-values for out-of-sample relative RMSE using different data vintages (Mark's sample period)
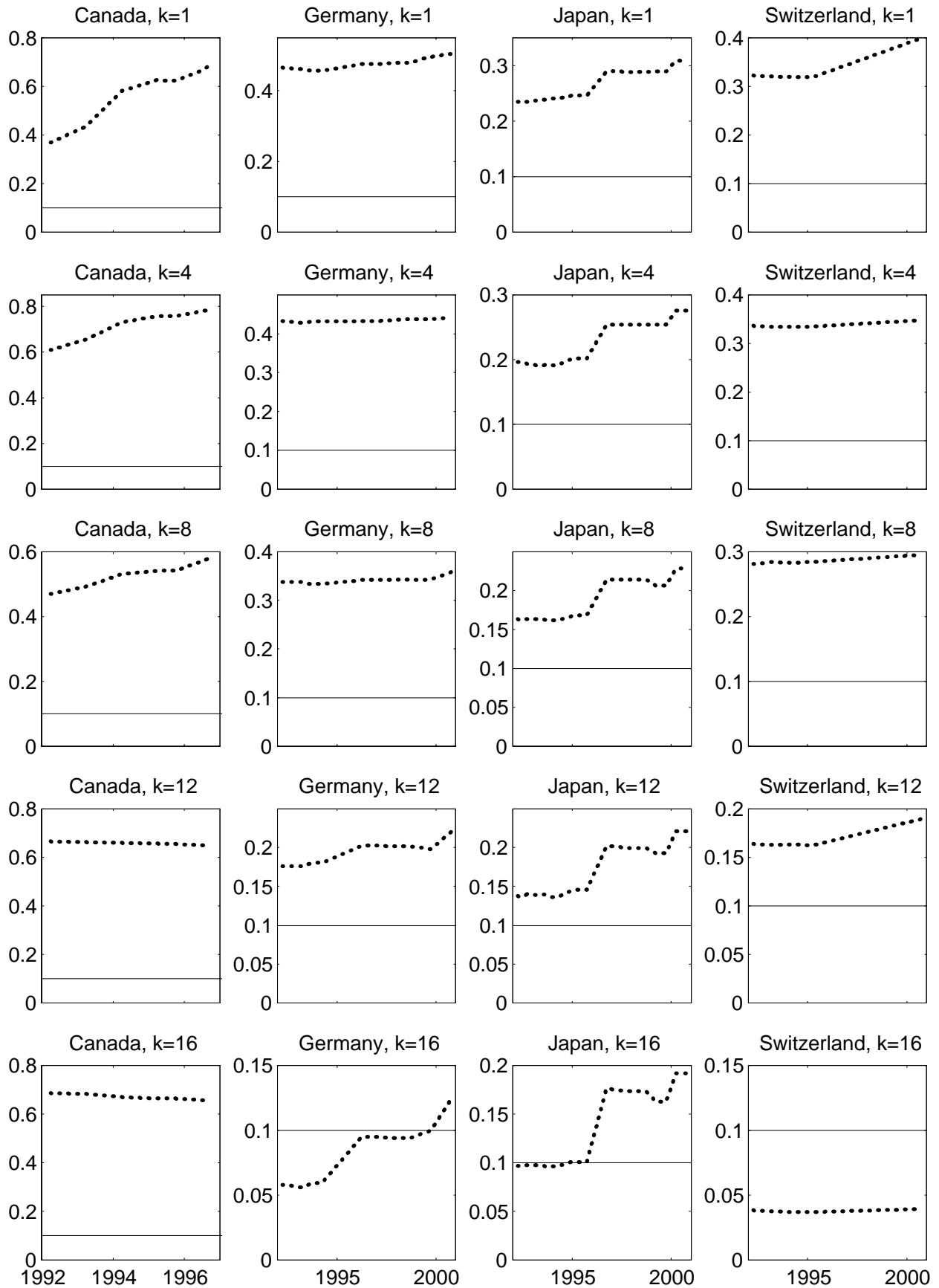
Figure 8: Estimates of β_k using different data vintages(Mark's sample period)
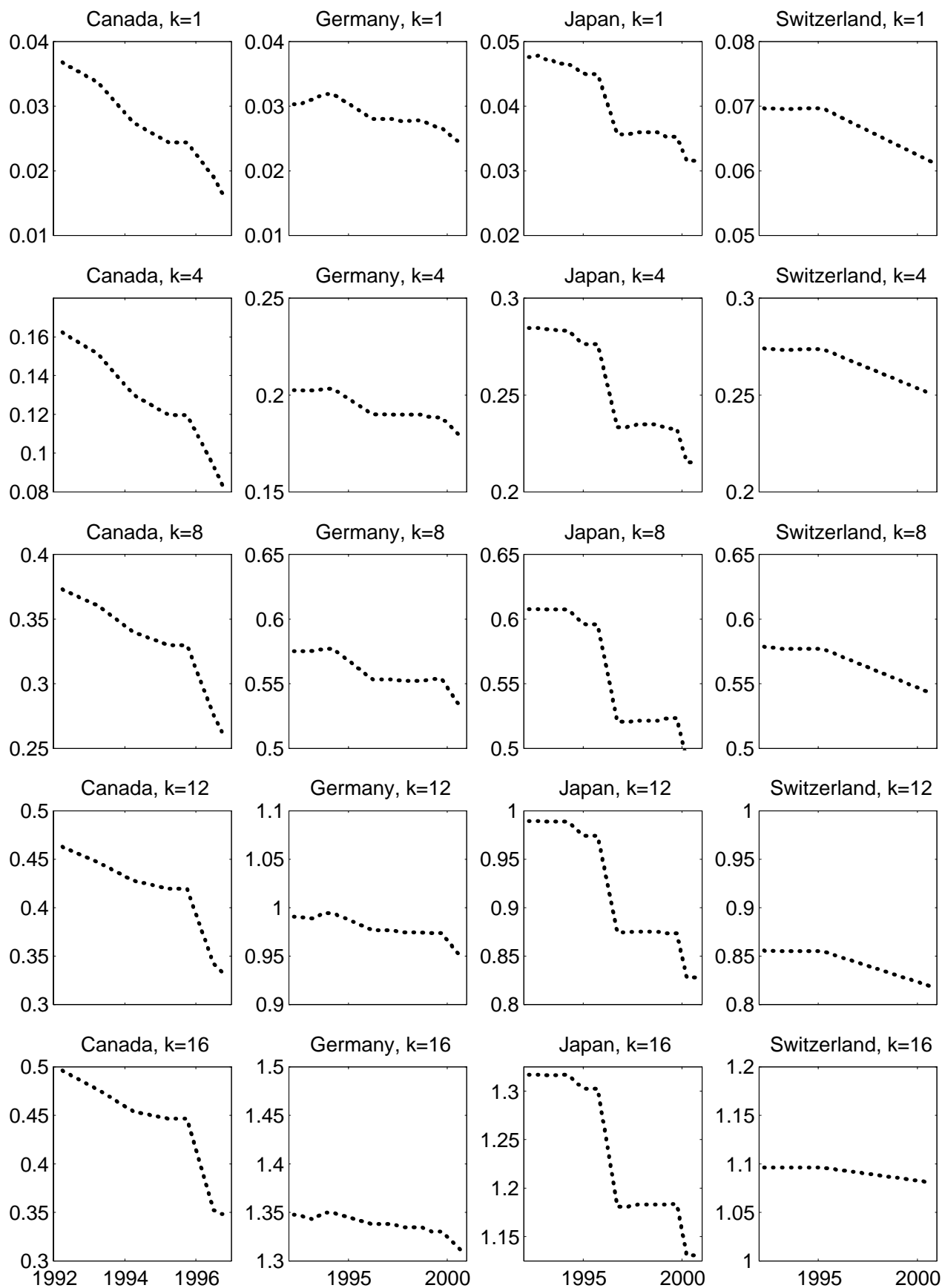
Figure 9: Bootstrap p-values for t_k using different data vintages (Mark's sample period)