Empirica
DOI 10.1007/s10663-015-9313-3

ORIGINAL PAPER

International stock return predictability: Is the role of U.S. time-varying?

Goodness C. Aye[1] · Mehmet Balcilar[1,2] · Rangan Gupta[1]

Abstract This study investigates the predictability of 11 industrialized stock returns with emphasis on the role of U.S. returns. Using monthly data spanning 1980:2–2014:12, we show that there exist multiple structural breaks and nonlinearities in the data. Therefore, we employ methods that are capable of accounting for these and at the same time date stamping the periods of causal relationship between the U.S. returns and those of the other countries. First we implement a subsample analysis which relies on the set of models, data set and sample range as in Rapach et al. (J Finance LXVIII(4):1633–1662, 2013). Our results show that while the U.S. returns played a strong predictive role based on the OLS pairwise Granger causality predictive regression and news-diffusion models, its role based on the adaptive elastic net model is weak. Second, we implement our preferred model: a bootstrap rolling window approach using our newly updated data on stock returns for each countries, and find that U.S. stock return has significant predictive ability for all the countries at certain sub-periods. Given these results, it would be misleading to rely on results based on constant-parameter linear models that assume that the relationship between the U.S. returns and those of other industrialized countries are permanent, since the relationship is, in fact, time-varying, and holds only at specific periods.

✉ Goodness C. Aye
goodness.aye@gmail.com

Mehmet Balcilar
Mehmet@mbalcilar.net

Rangan Gupta
rangan.gupta@up.ac.za

[1] Department of Economics, University of Pretoria, Pretoria 0002, South Africa

[2] Department of Economics, Eastern Mediterranean University, Via Mersin 10, Famagusta, Northern Cyprus, Turkey

Published online: 31 December 2015

Keywords Stock returns · Predictability · Structural breaks · Nonlinearity · Time varying causality

JEL Classification C32 · G10 · G15

1 Introduction

The recent financial and economic crisis has heightened research and policy attention to the stock market dynamics, in particular its predictability. This is because of the potential spill over effects from the stock markets to the real sector and the fact that they help in predicting output and inflation by acting as leading indicators (Stock and Watson 2003). Therefore, to design appropriate policies in advance for avoiding any impending crisis, there is need to predict stock returns accurately. There has been evidence of the U.S. and other international stock returns in-sample and out-of-sample predictability in a number of studies (Rapach and Wohar 2006; Ang and Bekaert 2007; Rapach and Zhou 2013; Henkel et al. 2011; Ferreira and Santa-Clara 2011; Dangl and Halling 2012; Gupta and Modise 2012; Rapach et al. 2013 e.t.c.). However, this has been questioned in few other studies (Bossaerts and Hillion 1999; Goyal and Santa-Clara 2003; Goyal and Welch 2008). Also the question of which variables have predictive ability is still an on-going debate. Common predictors in the literature include: valuation ratios (Campbell and Shiller 1998), the dividend yield (Rozeff 1984; Henkel et al. 2011; Rapach et al. 2013), the short interest rate (Ang and Bekaert 2007; Dangl and Halling 2012; Henkel et al. 2011; Rapach et al. 2013), the default premium (Fama and Bliss 1987; Campbell 1987; Fama and French 1989), the slope of the term structure (Keim and Stambaugh 1986; Campbell 1987; Fama and French 1989), long term yield and dividend-payout ratio (Dangl and Halling 2012; Gupta and Modise 2012), earnings growth (Ferreira and Santa-Clara 2011); price-dividend and price-earnings ratio (Ferreira and Santa-Clara 2011; Gupta and Modise 2012), debt ceiling and government shutdown (Aye et al. forthcoming) among others.

This study focuses on the lagged U.S. returns uncovered as a new predictor in Rapach et al. (2013). Using monthly data from 1980:2 to 2010:12 on 11 industrialized countries, Rapach et al. (2013) show that in many non-U.S. industrialized countries lagged U.S. returns significantly predict returns better than those countries' own economic variables, while lagged non-U.S. returns exhibit limited predictive power with respect to U.S. returns. Using news diffusion model, they show that U.S. return shocks are only fully reflected in equity prices outside of the U.S. with a lag. The economic rationale for including lagged U.S. returns as a predictor is based on the argument that returns in one country can predict returns in a trading-partner country if a two-country Lucas-tree framework with gradual information diffusion is employed (Hong et al. 2007; Rizova 2010). Therefore, given that U.S. has the largest equity market in the world in terms of market capitalisation, and is a trading partner for many countries, the market is likely to receive the most attention from investors, consequently causing a gradual diffusion of information on the global macroeconomic fundamentals from the U.S. market to other countries' markets (Rapach et al. 2013).

The current paper contributes to the international stock returns predictability literature by re-examining the in-sample predictive role of the lagged U.S. stock returns in a time varying framework. Specifically, we employ a bootstrap rolling window approach. Results in Rapach et al. (2013) are based on estimations from ordinary least squares (OLS), adaptive elastic nets and generalised method of moments (GMM) which are based on full samples. The use of full sample is based on the assumption that model parameters are constant over time. However, in an ever changing socioeconomic environment, this assumption may be quite restrictive. The assumption hardly ever holds and is a puzzling topic for economic empirical studies (Granger 1996). The presence of structural breaks and nonlinearities as is common with financial variables would therefore invalidate any conclusions from the full sample in-sample predictive estimations or the standard Granger causality results. A number of ways have been devised to account for structural breaks in economic relationships. The most common practice would be to test for the presence of structural breaks in advance and modify the estimation in various ways, for example, with the use of dummy variables or sample splitting. However, it has been argued that these methods can introduce some pre-test bias (Balcilar et al. 2010). This notwithstanding, we first perform subsample analyses using the same models in Rapach et al. (2013).

There are three main approaches, commonly employed in econometric applications, to estimation when structural breaks are likely: recursive, rolling, and time varying parameters (TVP). Recursive and TVP estimation are analogues as they keep the lower end of the estimation window and move towards and with a growing window. As the window size grows it accumulates more information and when they reach the last observation, they will be equivalent the full sample estimation (Inglesi-Lotz et al. 2014). If the parameters are stable, then the recursive and TVP estimators will converge to the constant parameters as the sample size grows with increasing window size. This implies that successive prediction errors will diminish for the estimate of the parameters, as the information already incorporated in the estimation increases. A consequence of this is that all previous observations will have impact on the successive estimates. In the presence of multiple structural breaks such an approach is not optimal since it will be difficult isolate the impact of previous breaks on the later ones. Armah and Swanson (2015) note that rolling window estimation is preferred by researchers due to lack of knowledge on the presence and number of structural breaks in the sample.

It may be a better idea to give more weight to more recent data when the parameters are not constant and it might even be more appropriate to discard the data that has reached certain age and passed the date of expiry. One way of better accommodating parameter variability is then to base the estimation only on the most recent portion of the data. This leads to what is know as rolling estimation. Our preference for rolling estimation is based on its better capability to accommodate parameter variations, particularly multiple ones.

Stock and Watson (1996) uses TVP and rolling estimation equivalently outperforms the other approaches. In application to time varying betas Groenewold and Fraser (1999) show that rolling estimation shows greatest variation in the sample and thus better captures the structural breaks. Barnett et al. (2012) also find that rolling estimation slightly outperforms other approaches, including TVP.

Moreover, Granger causality tests based on recursive estimation will have different sample sizes and the estimates will likely to be impacted by more number of structural breaks as the windows size grows. Therefore, the Granger causality tests based on recursive estimation are less comparable than the Granger causality tests based on rolling estimation. Rolling estimations might have higher parameter variance estimates as they are based on smaller sample sizes, however the precision of the estimates can be improved using bootstrap technique as it is used in this paper (see, Pesaran and Timmermann 1999; Clark and McCracken 2004). The impact of structural breaks is the main focus of this paper than the precision of the estimates, so the rolling estimation serves our purpose better.

Our bootstrap rolling window approach is robust to small samples and presence of multiple structural breaks and nonlinearities while also providing evidence of existence or otherwise of temporal causal relationship (in-sample predictability over time) between U.S. stock returns and international stock returns. Understanding the predictive role of the lagged U.S. returns has implications for asset international pricing models, hedging and investing behaviour and choices (Rapach et al. 2013).

2 Data and methodology

As earlier stated, we started by performing subsample analyses using OLS, adaptive elastic nets and GMM. For the estimation of these models we use data on excess stock returns, 3-month Treasury bill rates and dividend yield from 11 industrialized countries: Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland, United Kingdom and United States. These are the same data sets and sample range in Rapach et al. (2013) as we do not have access to all the series for all the countries. The summary statistics and original sources of these data are provided in Rapach et al. (2013).

However, for the rolling window estimation, we use updated excess stock returns from the 11 countries. We compute each country's equity premium (i.e. excess returns) as stock return less the annualized rate of the 3-month Treasury bill rate. The stock returns are computed as the first log differences of the stock prices of the relevant countries. The stock price and 3-month Treasury bill rates data are obtained from Thomson Reuters Datastream. The sample covers the period 1980:2–2014:12 after transformations with the exception of that of Sweden which span from 1982:3 to 2014:12. The starting and ending periods are determined by the availability of the 3-month Treasury bill rates data. We keep the stock returns in their respective national currencies to enable us analyse the predictive power of lagged U.S. returns for the other countries' returns.

As the OLS regressions, adaptive elastic nets and News-diffusion models only serve as a precursor to our preferred method, the bootstrap rolling window approach, we do not discuss them here.[1] So we turn to the bootstrap rolling window

[1] Interested readers may consult Rapach et al. (2013) for the details on these models.

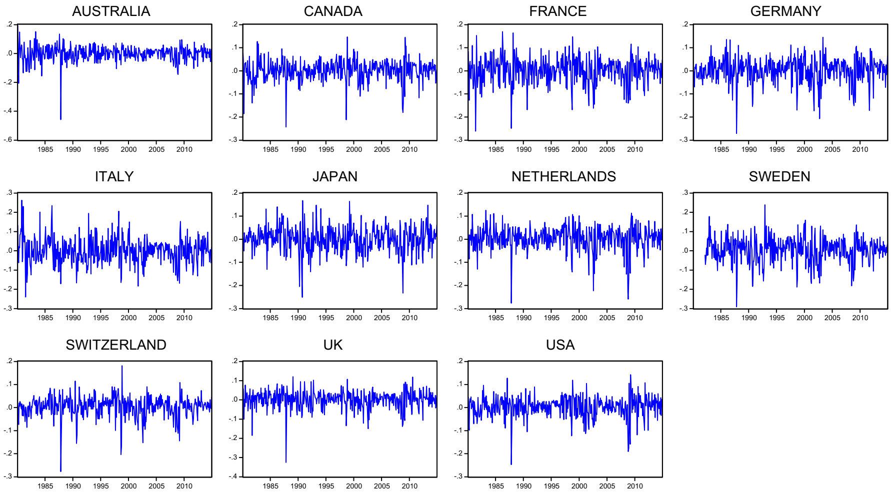

Fig. 1 Equity premium of the 11 industrialized countries

approach. Here the null hypothesis is Granger non-causality from U.S. returns to international returns. The assumption is that there is no causality (or predictive power) from international returns to U.S. returns, because of its large equity market concentration and a major trading partner for many countries. The joint parameter restriction associated with the Granger non-causality test in a VAR framework can be examined with the *Wald*, Likelihood ratio (*LR*) and Lagrange multiplier (*LM*) statistics based on the assumption that the underlying series is stationary, which is the case in this study, given the nature of the data transformation (see Fig. 1). Hence, we do not use the Toda and Yamamoto (1995) procedure for testing for Granger causality.

Building on the standard Granger non-causality test, we use a residual based (*RB*) bootstrap test rather than standard asymptotic tests while accounting for the fact that international returns has no in-sample predictability for U.S. returns. Following Balcilar and Ozdemir (2013) and Balcilar et al. (2010), we use the *RB* based modified-*LR* statistics to examine the causality between U.S. returns and international returns.

The bootstrap modified-*LR* Granger causality can be demonstrated starting with a bivariate VAR(p) process of the form:

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a white noise process with zero mean and covariance matrix Σ and p is the lag order of the process. We use the Schwarz Information Criterion (SIC) to select the optimal lag order p in the empirical section. For simplification, let y be partitioned into two sub-vectors, y_1 (U.S. returns) and y_2 (international returns). Hence, Eq. (1) can be rewritten in a matrix format as follows:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varphi_{10} \\ \varphi_{20} \end{bmatrix} + \begin{bmatrix} \varphi_{11}(L) & 0 \\ \varphi_{21}(L) & \varphi_{22}(L) \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{2}$$

where $\phi_{ij}(L) = \sum_{k=1}^{p+1} \phi_{ij,k} L^k, \quad i,j, = 1,2$ and L is the lag operator such that $L^k y_{it} = y_{it-k}, \quad i = 1,2$. The restriction $\varphi_{12}(L) = 0$ in Eq. (2) is due to the exogeneity assumption of the U.S. stock returns. We test the null hypothesis that U.S. returns does not Granger cause international by imposing zero restrictions: $\varphi_{21,i} = 0$ for $i = 1, 2, \ldots, p$. This implies that if the joint zero restrictions under the null hypothesis:

$$H_0 : \varphi_{21,1} = \varphi_{22,2} = \cdots = \varphi_{21,p} = 0. \tag{3}$$

are not rejected, then U.S. returns does not cause or contain predictive ability for international returns. If the hypothesis in Eq. (3) is rejected, then U.S. returns Granger causes international returns. The causality hypothesis in Eq. (3) can be tested using a number of testing techniques. However, this study uses the bootstrap approach which uses critical or p-values generated from the empirical distribution derived for the particular test using the sample data.

Structural changes shift the parameters and the pattern of the causal relationship may change over time. To deal with structural changes and parameter non-constancy, this paper applies the bootstrap causality test to rolling window subsamples for $t = \tau - l + 1,\ \tau - l, \ldots, \tau,\ \tau = l,\ l + 1, \ldots, T$, where l is the size of the rolling window.[2] We apply the causality test to each subsample in each step, providing a (T − 1) sequence of causality tests instead of only one. This also allows us to detect whether U.S. returns has led international returns over time. We test for the existence of structural breaks using Bai and Perron (2003) tests for multiple structural.

3 Results

3.1 Preliminary analysis

The plots of the monthly country excess stock returns are shown in Fig. 1. These show that all the series are stationary and this is not surprising given that the stock returns upon which excess stock returns are calculated are the differenced first natural logs of stock prices. We also present the summary statistics for the monthly excess returns (in percent) for the 11 countries in Table 1. During the sample period, Sweden has the highest average returns (0.40 %) followed by Switzerland and U.S. while Australia has the least (−0.10 %). Italy displays the greatest volatility over the sample period. All countries have positive autocorrelation with Switzerland displaying the largest value (0.18) while U.K. display the smallest autocorrelation (0.02).

Prior to estimating the relevant models in Rapach et al. (2013), we perform the Bai and Perron (2003) multiple break test. For the U.S. equity premium, the Bai–

[2] More technical details on the approach we use can be found in Balcilar et al. (2010).

Table 1 Summary statistics of monthly country excess stock returns

Country	Mean	SD	Minimum	Maximum	Autocorrelation
Australia	−0.10	5.35	−45.82	19.27	0.08
Canada	0.07	4.60	−24.26	14.69	0.12
France	0.15	5.79	−26.10	16.93	0.12
Germany	0.16	5.43	−27.16	14.55	0.11
Italy	0.01	6.82	−24.00	26.26	0.12
Japan	0.09	5.55	−25.12	16.74	0.12
Netherlands	0.23	5.31	−27.66	12.50	0.12
Sweden	0.40	6.63	−29.07	23.88	0.16
Switzerland	0.39	4.56	−27.78	18.11	0.18
United Kingdom	0.09	4.72	−32.57	12.11	0.02
United States	0.29	4.56	−24.72	14.30	0.06

Perron test is performed by regressing the equity premium of the U.S. on a constant only, and is reported in Table 3. We find evidence of five significant break points in the U.S. equity premium. Based on these we extract 3 subsamples (1982:09–2000:08 and 2000:09–2010:12; 1982:09–2002:09 and 2002:10–2010:12; 1982:09–2007:5 and 2007:6–2010:12) and perform subsample analysis instead of the full sample (1980:2–2010:12) analysis as in Rapach et al. (2013).

Further, the condition for using the rolling window causality testing approach also depends on the evidence of instability in the relationships. Therefore, we also test for the presence of structural breaks. We regress country *i* returns on a constant, one lag of U.S. returns and one lag of country *i* returns. The results are also presented in Table 3. In most cases, we observe as many as five significant breaks. With the presence of structural breaks, the assumption of constant parameters over time as in full sample predictive regressions or standard Granger causality tests is no longer valid. Hence, we proceed with the subsample analysis and rolling window regression approach.

3.2 Subsample results

The OLS estimates of the benchmark predictive regression of national 3-month Treasury bill rate $(\beta_{i,b})$ and log dividend yield $(\beta_{i,d})$ on equity premium for each country *i* are reported in Table 4. The estimates are reported in columns 1, 2, 4, 5, 7 and 8 with their corresponding *t*-statistics (based on heteroskedasticity-robust standard errors) in parentheses.[3] We also report the corresponding R^2 statistics. The values in parentheses under the R^2 statistics are the χ^2 statistics for testing the null hypothesis that $\beta_{i,b} = \beta_{i,d} = 0$, implying no return predictability for country *i*. For brevity, we highlight the coefficient estimates and R^2 statistics that are significant at

[3] To conserve space we do not report the wild bootstrapped *p*-values here but these are available from authors upon request.

the 10 % level or better in bold fonts. Overall, based on the wild bootstrapped p-values, we observe a more predictive ability of the nominal interest rate than the dividend yield consistent with findings in Rapach et al. (2013). However, in contrasts to their results, we obtain more robust estimates and significant results for both variables and for more countries. For instance, while dividend yield was a significant return predictor for only U.K. in Rapach et al. (2013), here it is significant return predictor for at least five countries in all the subsamples with exception of the most recent periods. With respect to the R^2 statistics, a value near 0.5 % indicates economically significant return predictor (Kandel and Stambaugh 1996; Campbell and Thompson 2008). Our results show evidence of R^2 statistics above 1 % in general. Our R^2 statistics are also far larger than those reported in Rapach et al. (2013). For example while their largest R^2 is 2.6 % for U.K., we have 14.99 % for U.K. and 27.19 % (the largest in our results) for Sweden in the 2007:6–2010:12 sub-period. Also the null hypothesis of no return predictability indicated by the χ^2 statistics are rejected for more countries in each subsample analysis than in Rapach et al. (2013). A pooled version of the predictive regression which imposes $\beta_{i,b} = \bar{\beta}_b$ and $\beta_{i,d} = \bar{\beta}_d$ for all i while allowing for country-specific constants as reported in the last but two rows also produced a completely different result with respect to the size and significance of the coefficients, R^2 and χ^2 statistics. No significant results on these were found in Rapach et al. (2013). The signs here are however consistent with theirs with negative and positive coefficients for nominal interest rate and dividend yield respectively. These findings are not surprising given that we account for structural breaks in our analysis.

In Table 5, we report the results on the lead-lag relationship i.e. the pairwise Granger causality between country i and country j returns. These are obtained from a specification that allows us to include lagged country i and lagged country j excess returns as predictors of country i returns while controlling for predictive ability of national economic variables using the nominal interest rate and dividend yield. With exception of Japan and Switzerland, U.S. returns exhibit significant predictive power for all other countries returns at one sub-period or the other. It significantly predicts returns in 34 out of 66 cases (including the pooled version) lagging slightly behind Sweden with 35 significant coefficients and has the largest coefficient in 15 cases following Switzerland with 17 cases. However, only Swedish returns out of the 10 non-U.S. returns consistently shows predictive ability for U.S. returns except in the last sub-period (i.e. 5 out of 6 cases) while Switzerland and Australia show significant predictive power for the U.S. returns once. Moreover relatively large values for U.S. returns in the last column compared to its values on the last but one row is an indication of U.S. leading role in the international equity market consistent with Rapach et al. (2013). These results may justify our exogeneity assumption for the U.S. returns in the rolling window estimations to be discussed later. We note however that based on the average estimates and pooled model results, U.S. and Switzerland appear to be competing with each other.

Next as in Rapach et al. (2013), we estimate the adaptive elastic net (Zou and Zhang 2009; Ghosh 2011) model meant to improve the power of the test and precision of estimates. This specification allows us to control for all other country returns when testing for causality. Finally on the subsample analysis, we estimate a news-diffusion model that allows for a return shock from one country to be fully incorporated into another country with a lag, thereby permitting cross-country information frictions (Rapach et al. 2013). Overall while we find strong evidence of international spillover effects based on the news-diffusion models, the evidence based on the adaptive elastic net model is weak in contrast to Rapach et al. (2013).[4]

3.3 Results based on time varying bootstrap rolling causality

In this section we present the results from the bootstrap rolling window results as this is not only capable of handling multiple structural breaks but also accounts for nonlinearities in the causal relationships as well as robust to small sample sizes.

Two important decisions that must be made prior to estimation of the rolling window approach are the window size and lag order selection. With respect to window size, there is no strict selection criterion; however there is a trade-off between the accuracy of the parameter estimates and the representativeness of the model over the subsample period. On one hand, a small window size reduces heterogeneity and improves the representativeness of parameters, but it may reduce parameter accuracy by increasing the standard errors of estimates. On the other hand, a large window size may improve the accuracy of estimates, but reduces the representativeness of the model, especially in the presence of heterogeneity. Through Monte Carlo simulations, Pesaran and Timmermann (2005) showed that the bias in autoregressive (AR) parameters is minimized with window size around 10–20 when there are frequent breaks as in our case. Therefore we use a rolling window of small size of 24 months and apply bootstrap technique to each window so as to estimate the tests with better precision.[5]

The rolling window method uses a fixed length moving window sequentially from the beginning to the end of the sample by adding one observation from ahead and dropping one from behind, where each rolling window subsample includes l observations. To investigate potential changes in the causality relationships, we estimate the bootstrap value of observed LR-statistics rolling over the whole sample period 1980:2–2014:12, except Sweden for which the sample period is from 1982:3 to 2014:12. That is, we estimate the VAR model in Eq. (1) for a time span of 24 months rolling through $t = \tau - 23, \tau - 22, \ldots, \tau$, $\tau = 24, \ldots, T$. The bootstrap LR-test uses the p-values obtained from 1000 replications. We use the optimal lag in each of the windows for the VAR model estimations as determined by both SIC and AIC criteria.

The bootstrap p-values pertaining to the null hypothesis that U.S. returns does not have predictive power over country i returns are presented in Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11. Non-causality in each rolling subsample estimate is evaluated at a

[4] The results of the adaptive elastic net and news-diffusion models are available from the authors upon request.

[5] Notwithstanding, we estimated the model using longer window sizes (30 and 60), but find no qualitative changes in our results. Hence, we report here results based on a window of 24 months.

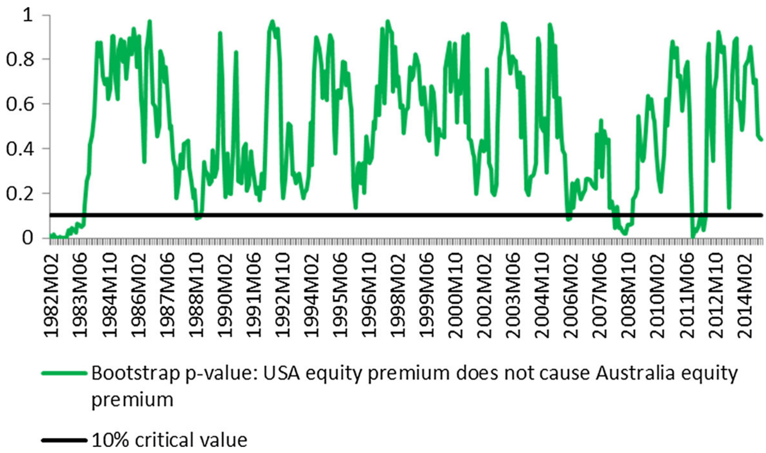

Fig. 2 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Australia equity premium

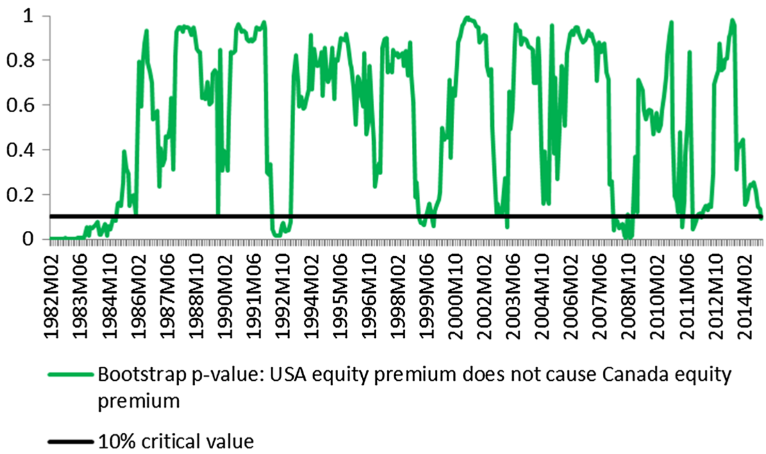

Fig. 3 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Canada equity premium

10 % level to guard against the low power of the test. It can be observed that in all cases, the *p*-values change substantially over the sample. The figures show the bootstrap *p*-values of the rolling test statistics testing the null hypothesis that the U.S. returns does not Granger cause or have predictive power for each country's returns. The null hypothesis is rejected at 10 % significance level for a number of sub-periods in each of the country as indicated when the *p*-values lie below the 10 % critical value. These significant periods of spillover effect of the US equity premium are further summarized in Table 2 for each country.

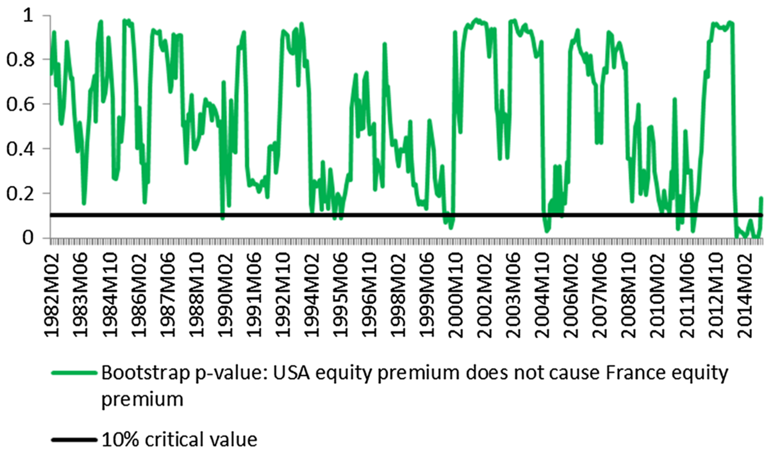

Fig. 4 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause France equity premium

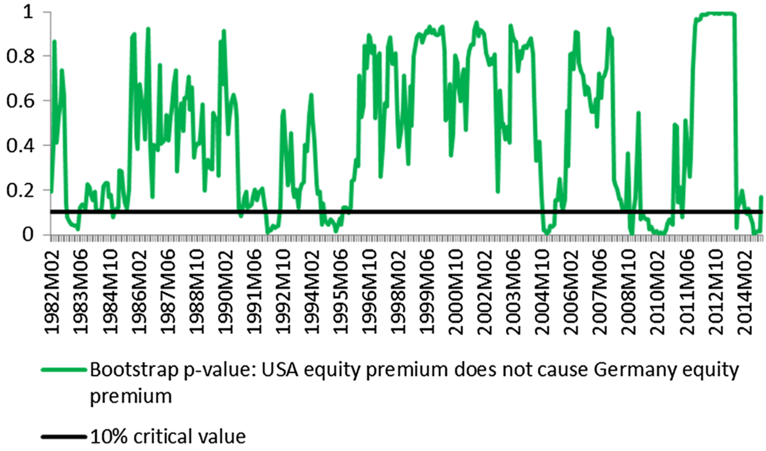

Fig. 5 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Germany equity premium

4 Conclusion

This study analyse the lead-lag relationship among 11 industrialized country (Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland, United Kingdom and United States) stock returns with specific aim of identifying the predictive role of the U.S. return. We use the national economic variables (dividend yield and 3-month Treasury bill rates) as control variables. Our data is monthly data covering the 1980:2–2014 period for all countries except Sweden for which data is available only from 1982:3 to 2014:12. Although, the idea behind this study is based on Rapach et al. (2013), we contribute by accounting for

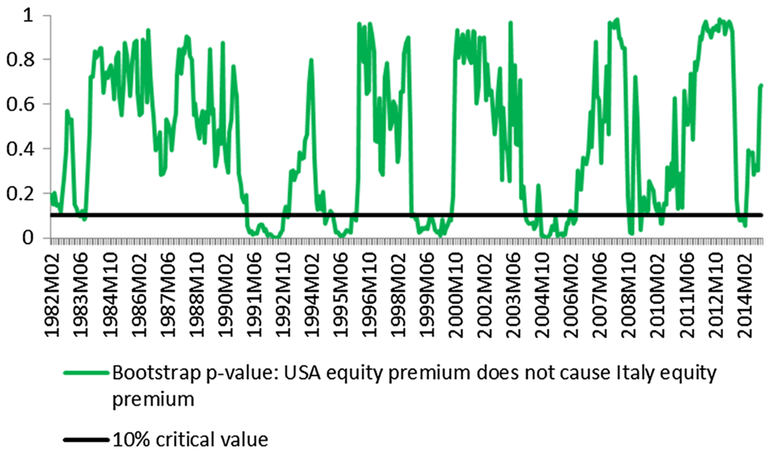

Fig. 6 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Italy equity premium

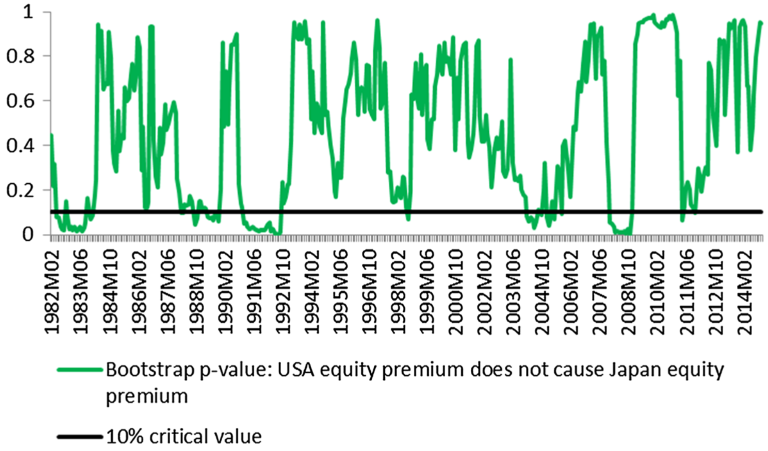

Fig. 7 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Japan equity premium

structural breaks and nonlinearities that pose challenge to financial time series data since these properties invalidate results from full sample standard Granger causality tests. This we do by employing two different approaches: a subsample analysis that are based on the same set of models (OLS, adaptive elastic net and news-diffusion models) estimated in Rapach et al. (2013) and a bootstrap rolling window causality test. The rolling window approach does not only account for multiple structural breaks, it is capable of dating exactly the periods for which the U.S. returns has predictive power for the international returns and it is robust to small sample size.

To determine the suitability of these two approaches, we first conduct multiple structural breaks and linear dependency tests. We find the existence of multiple

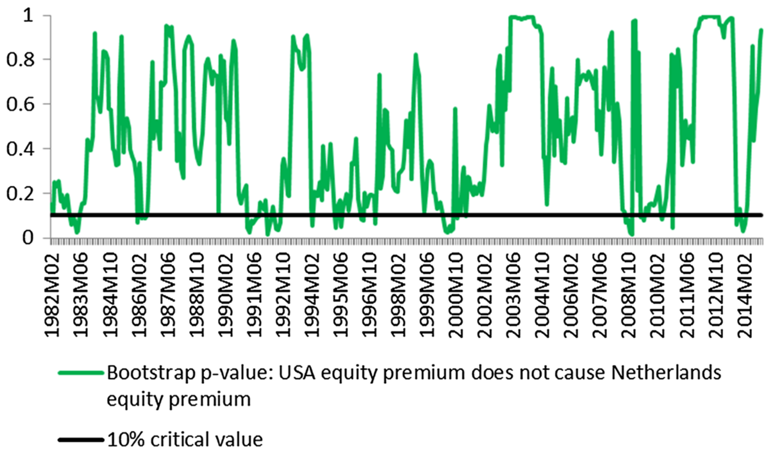

Fig. 8 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Netherlands equity premium

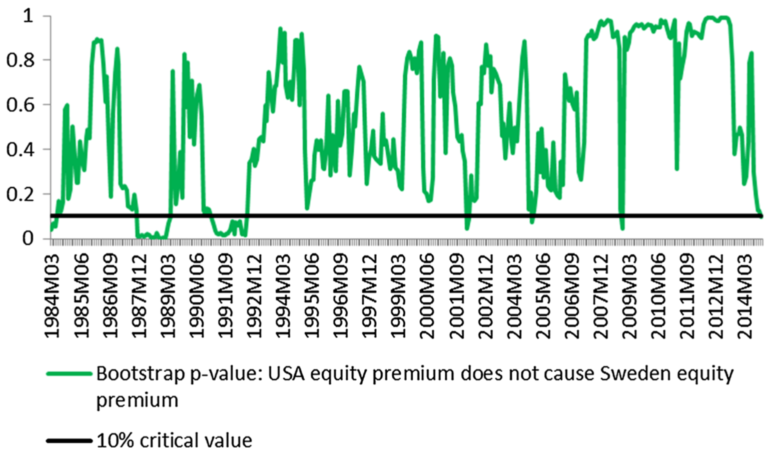

Fig. 9 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Sweden equity premium

structural breaks and nonlinearities in the data. Given this outcome we proceed first with the subsample analysis using the same data set and sample range as in Rapach et al. (2013). The subsample results based on the pairwise Granger causality predictive regression and the News-diffusion model in general support the findings in Rapach et al. (2013): the lagged U.S. returns has predictive power over other countries returns and that information friction plays a key role in the impact of U.S. return shocks on other countries. However, in contrast to Rapach et al. (2013) we do not find much evidence of the U.S. returns predictive power when adaptive elastic net models. Also we obtain more robust estimates in almost all cases than they did.

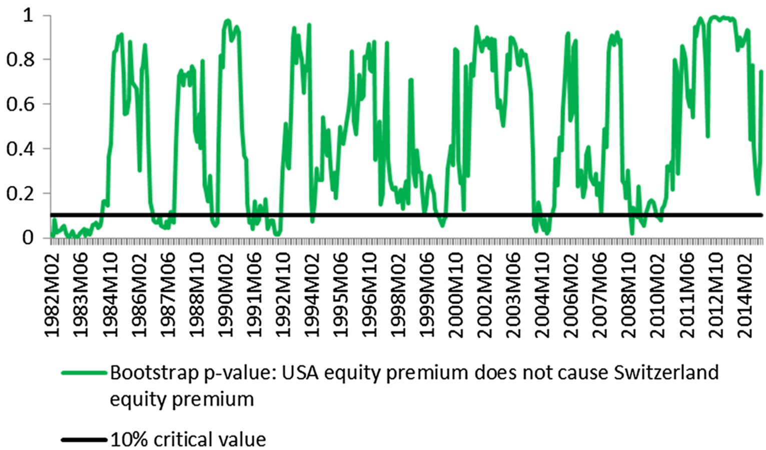

Fig. 10 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause Switzerland equity premium

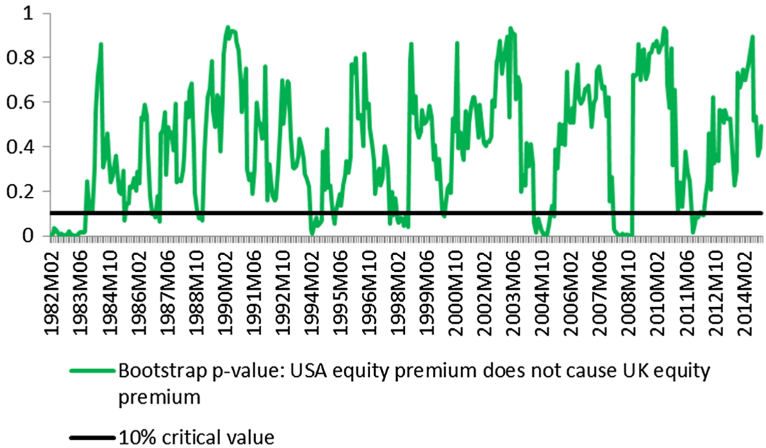

Fig. 11 Rolling window bootstrap *p*-value: USA equity premium does not Granger cause UK equity premium

It is also important to note that in all the estimations, the results vary from one subperiod to the other both in terms of size and significance.

Based on our new updated data set and the bootstrap rolling window approach, we find that there are certain subperiods during which significant spillover effects can be observed. In other words we find evidence that the causal relationship between U.S. returns and international returns vary over time. These periods differ across countries, and alternate with long periods during which there are no signs of a spillover effect. Hence, the results reported by Rapach et al. need some qualification. The time variation in the causal relationship between the U.S. returns and international returns invalidates any results based on the linear models since these assume a permanent relationship. Hence, this feature needs to be taken into

Table 2 Significant causal sub-periods

Country	Significant periods of causality of US equity premium
Australia	1982:2–1983:8, 1988:11–1989:1, 2006:1–2006:2, 2008:3, 2008:5–2008:12 and 2011:10–2012:5
Canada	1982:2–1985:2, 1992:5–1993:3, 1999:2–1996, 1999:9–1999:10, 2003:3, 2008:2–2008:12, 2014:4, 2011:10–2011:12, 2012:3 and 2014:12
France	1995:3, 1995:7, 2005:5, 2000:7–2000:9, 2004:12–2005:2, 2005:5, 2005:9, 2011:2, 2011:4, 2011:10, 2013:10–2014:11
Germany	1982:11–1983:5, 1984:1–1984:6, 1984:12–1985:1, 1990:11–1991:12, 1992:2–1992:8, 1994:10–1995:7, 1995:11, 2004:11–2005:5, 2008:9, 2008:11–2009:1, 2009:5–2010:11, 2011:10, 2013:10–2014:11
Italy	1983:6, 1983:8–1983:9, 1991:3–1992:11, 1993:1, 1994:10–1994:12, 1995:3–1996:3, 1998:10–2000:8, 2004:1–2004:7, 2004:10–2006:1, 2006:4–2006:5, 2008:11–2008:12, 2009:5, 2010:4–2010:5 and 2013:12–2014:3
Japan	1982:5–1982:9, 1982:11–1983:9, 1983:11–1984:1, 1988:9–1988:11, 1989:2, 1989:5–1989:11, 1991:1–1992:9, 1998:8, 2004:3–2004:10, 2005:1–2005:3, 2005:5, 2007:12–2008:12, 2011:4 and 2011:11
Netherlands	1982:12–1983:5, 1986:4–1986:1986:7, 1991:3–1991:9, 1992:2–1992:4, 1992:6–1992:9, 1994:3, 1995:3–1995:4, 1995:7, 1996:6–1996:7, 1997:2, 2000:4–2000:9, 2001:1, 2001:4, 2008:9–2008:12, 2009:5–2009:7, 2009:9, 2010:4–2010:5, 2010:11, 2013:10–2013:11 and 2014:1–2014:3
Sweden	1984:3–1984:5, 1987:12–1989:5, 1990:11, 1991:2–1992:8, 2005:1 and 2008:12–2014:12
Switzerland	1982:2–1984:6, 1986:11–1987:7, 1987:9–1987:10, 1989:7–1989:10, 1991:4–1991:6, 1991:8–1991:9, 1991:11–1991:12, 1992:2–1992:9, 1994:3, 1995:7, 1999:12–2000:5, 2004:7, 2004:10–2005:4, 2007:7, 2008:12, 2009:3, 2009:5–2009:7 and 2010:2–2010:4
UK	1982:2–1983:9, 1985:7, 1986:10, 1986:10, 1986:12, 1987:2, 1988:12–1989:2, 1994:3–1994:8, 1995:2–1995:4, 1997:10, 2008:2–2008:12 and 2011:10–2012:4

account when modeling and predicting stock returns. The fact that international returns are predictable is interesting given that they act as leading indicators in the economy and hence serve as a source of useful information for policy makers and investors as to where the economy might be heading. Given that US is the world's largest in terms of GDP and a major trading partner to most countries shocks to the US stock market could be transmitted to these other markets albeit not at every point in time. A policy action may be required by these other countries in the event of negative shocks to the US stock market. For example capital controls could delay the reaction of a country's stock market to news about the US stock markets. The results also show the need to explicitly incorporate the US leading role in building international asset pricing models. We could not control for fundamentals by using same national economic variables as in Rapach et al. (2013) in the rolling window estimations due to data unavailability for all countries, therefore we suggest that future studies should incorporate these once they become available.

It should also be noted that the economic developments that may have led to observed spillover effect have not been the subject of our empirical research, but, future research could try to recover whether the spillover episodes are associated with some of key market, technological and regulatory events in the U.S. stock market Key market events that could be investigated include the 1987 Black Monday stock market crash, 13 Oct 1989 mini-crash caused by failed leveraged buyout of United Airlines, early 1990s recession caused by invasion of Kuwait by Iraq in July 1990, the 1997–2000 dot-com bubble burst, 2001 September 11 attacks, the stock market downturn of 2002, 2007 Quant crash, United States bear market of 2007–09, 2007–2008 subprime credit crisis periods and the 2010 Flash Crash Perhaps the 2012 Presidential election, September 13, 2012 Federal Reserve announcement of a third round of quantitative easing (QE3) and the continued debate on Fiscal Cliff can also be investigated. The regulatory events worth investigating include: the 1996 NASDAQ litigation, 1997 Order handling rules that prompted rise of Electronic Communication Networks (ECNs) for transaction costs reduction, the 1999 Regulation of Alternative Trading Systems (ATS) which allows ECNS to operate as broker dealers without exchange registration and eliminates any market making obligations and 9 April 2001 Decimalization that facilitated smaller lots and market automation, the 2005 Regulation of the National Market System (NMS II), the 2008 Short Sale Bans and the 2011 Uptick Rule which restricts short selling. With respect to technological events, the following could be investigated: the 1980s program trading, that is the simultaneous trading of a portfolio of stocks, as opposed to buying or selling just one stock at a time, 1988 Small Order Execution System (SOES 1999 Instinet Order Management System (OMS) first Execution management systems (EMS) platform and 2001 Credit Suisse (CS) Advanced Execution Services (AES) launch.

Acknowledgments We wish to thank the referee for the helpful comments. However, any remaining errors are solely ours.

Appendix

See Tables 3, 4, 5, 6 and 7.

Table 3 Multiple structural break tests

Country	No. of breaks	Estimated break dates				
		1st	2nd	3rd	4th	5th
Australia	5	1982M08	1987M11	1990M12	2007M05	2009M01
Canada	5	1982M08	1998M11	2000M10	2002M12	2009M01
France	5	1998M11	2000M10	2002M11	2007M07	2009M04
Germany	5	1987M01	1992M12	1998M03	2003M05	2009M01
Italy	5	1984M06	1986M05	1990M08	1992M11	1994M07
Japan	5	1990M08	1992M12	2002M07	2008M12	2012M07
Netherlands	2	2001M08	2003M07	–	–	–
Sweden	5	1989M09	1992M12	2000M04	2002M11	2006M06
Switzerland	3	1996M09	1998M08	2000M04	–	–
United Kingdom	5	1982M08	1987M09	1990M12	2001M07	2009M01

Table 4 Benchmark predictive regression model results

Country (*i*)	1982:9–2000:8			2000:9–2010:12			1982:9–2002:9		
	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$
Australia	−0.03	2.03	0.44	**−0.96**	−1.50	5.76	−0.02	2.14	0.58
	(−0.2)	(0.47)	(0.41)	(−1.98)	(−0.43)	(5.01)	(−0.11)	(0.52)	(0.87)
Canada	**−0.32**	1.65	1.73	**−0.74**	−0.46	4.63	**−0.39**	**3.64**	2.94
	(−1.73)	(0.91)	(3.76)	(−2.65)	(−0.24)	(7.2)	(−2.17)	(2.12)	(4.99)
France	**−0.43**	**4.40**	**2.46**	**−1.42**	−1.75	**8.61**	−0.19	**3.38**	1.43
	(−2.36)	(2.45)	(6.43)	(−2.72)	(−0.6)	(10.81)	(−0.99)	(1.8)	(4.08)
Germany	−0.47	2.60	1.79	**−1.67**	−3.24	**9.38**	−0.30	3.03	1.29
	(−1.68)	(1.31)	(2.87)	(−3.97)	(−1.11)	(15.79)	(−1.05)	(1.55)	(2.4)
Italy	−0.05	−0.34	0.13	**−1.18**	−1.79	**6.27**	0.12	−1.93	0.82
	(−0.32)	(−0.18)	(0.25)	(−2.67)	(−0.7)	(7.53)	(1.09)	(−1.14)	(1.91)
Japan	0.04	**1.90**	**1.44**	**−5.63**	1.06	**5.76**	0.16	**1.55**	0.92
	(0.2)	(2.28)	(5.25)	(−2.7)	(0.63)	(8)	(0.92)	(1.93)	(4.17)
Netherlands	**−0.66**	**2.98**	**4.62**	**−1.69**	−2.14	**12.53**	**−0.47**	**3.05**	**2.91**
	(−4.06)	(2.35)	(16.53)	(−3.66)	(−0.87)	(15.08)	(−2.51)	(2.53)	(8.02)
Sweden	**−0.40**	**4.51**	**2.64**	**−1.74**	1.23	**14.41**	−0.04	2.12	0.57
	(−2.49)	(1.8)	(6.27)	(−3.87)	(0.52)	(17.85)	(−0.25)	(0.88)	(0.89)
Switzerland	**−0.43**	1.90	**2.49**	**−1.01**	−1.47	**4.62**	**−0.33**	**2.30**	**1.67**
	(−2.59)	(1.59)	(7.08)	(−2.69)	(−0.74)	(7.3)	(−2.02)	(2.01)	(5.83)
United Kingdom	**−0.23**	**3.99**	2.47	**−0.54**	0.45	**4.95**	**−0.17**	**4.75**	**4.11**
	(−1.57)	(2.12)	(4.55)	(−1.77)	(0.13)	(6.06)	(−1.16)	(3.04)	(10.89)
United States	−0.10	0.53	0.18	−0.15	2.45	2.26	−0.02	1.53	2.07
	(−0.58)	(0.57)	(0.42)	(−0.72)	(0.84)	(1.83)	(−0.11)	(1.79)	(4.59)
Pooled	**−0.17**	**1.47**	**1.06**	**−1.03**	−0.82	**6.09**	−0.03	**1.45**	**0.71**
	(−2.09)	(2.28)	(7.12)	(−4.82)	(−0.49)	(23.69)	(−0.41)	(2.52)	(6.55)

Table 4 continued

Country (i)	2002:10–2010:12			1982:9–2007:5			2007:6–2010:12		
	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$	$\hat{\beta}_{i,b}$	$\hat{\beta}_{i,d}$	$R^2(\%)$
Australia	**−1.33**	−5.08	**9.15**	−0.05	2.43	0.58	**−1.56**	−3.21	**16.81**
	(−2.14)	(−1.1)	(5.26)	(−0.37)	(0.62)	(0.53)	(−2.27)	(−0.62)	(7.4)
Canada	**−0.50**	−3.35	4.31	**−0.42**	**3.67**	**3.55**	**−1.26**	−7.02	**10.34**
	(−1.79)	(−1.59)	(4.99)	(−2.71)	(2.25)	(7.42)	(−2.7)	(−1.11)	(7.32)
France	**−1.28**	−3.13	**7.19**	−0.16	**3.00**	**1.64**	**−1.35**	1.43	**14.64**
	(−2.42)	(−0.99)	(5.96)	(−1.57)	(2.37)	(5.66)	(−2.37)	(0.28)	(7.93)
Germany	**−1.43**	−3.45	**8.07**	−0.44	2.88	1.50	**−1.82**	−2.42	**17.76**
	(−2.94)	(−1.13)	(8.7)	(−1.68)	(1.55)	(2.92)	(−3.3)	(−0.43)	(11.46)
Italy	**−0.99**	−2.79	5.82	−0.01	−0.14	0.01	**−1.39**	1.22	**11.98**
	(−2.11)	(−0.92)	(4.75)	(−0.07)	(−0.11)	(0.01)	(−2.69)	(0.29)	(7.38)
Japan	**−6.90**	−0.10	**9.79**	0.11	**1.82**	**1.18**	−5.87	5.32	**11.88**
	(−3.13)	(−0.06)	(9.77)	(0.71)	(2.38)	(5.65)	(−0.93)	(1.01)	(9.38)
Netherlands	**−1.54**	−2.89	**13.70**	**−0.36**	**2.61**	**2.37**	**−1.73**	0.01	**19.96**
	(−2.95)	(−1.02)	(8.91)	(−2.59)	(2.31)	(7.86)	(−2.95)	(0)	(8.9)
Sweden	**−1.42**	−1.27	**11.05**	−0.08	2.83	1.32	**−2.08**	2.42	**27.19**
	(−2.71)	(−0.44)	(7.88)	(−0.93)	(1.73)	(3.88)	(−3.55)	(0.66)	(12.85)
Switzerland	−0.91	−2.47	3.53	**−0.36**	**2.62**	**2.25**	−1.26	−0.49	5.90
	(−1.79)	(−1.25)	(4.16)	(−2.58)	(2.24)	(7.83)	(−1.44)	(−0.08)	(3.45)
United Kingdom	**−0.61**	−3.38	5.35	**−0.23**	**4.88**	**3.89**	**−0.98**	−1.30	**14.99**
	(−1.93)	(−0.78)	(3.99)	(−1.76)	(3.24)	(11.3)	(−2.68)	(−0.29)	(7.77)
United States	−0.21	−1.36	0.60	−0.14	1.65	1.69	−0.90	−1.08	3.94
	(−1.02)	(−0.33)	(1.04)	(−0.97)	(2.01)	(4.35)	(−1.59)	(−0.16)	(3.44)
Pooled	**−0.90**	−2.66	**5.81**	**−0.08**	**1.56**	**0.89**	**−1.38**	0.06	**13.20**
	(−3.47)	(−1.31)	(15.03)	(−1.28)	(3.05)	(9.71)	(−3.13)	(0.02)	(10.03)

Table 5 Pairwise Granger causality test for first subsample: 1982:9–2000:8 and 2000:9–2010:12

Country (i)	1982:9–2000:8										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		0.09	**0.15**	**0.17**	**0.08**	**0.10**	0.11	**0.08**	0.05	0.06	0.12
		(0.85)	(1.78)	(2.04)	(1.82)	(1.58)	(1.1)	(1.64)	(0.64)	(0.62)	(1.12)
CAN	0.00		0.02	−0.04	0.00	0.00	−0.10	0.07	−0.06	0.00	0.02
	(−0.03)		(0.24)	(−0.6)	(0.04)	(−0.03)	(−1.2)	(1.59)	(−0.69)	(−0.05)	(0.2)
FRA	0.01	0.04		−0.08	0.00	−0.01	−0.07	**0.12**	0.03	0.05	0.11
	(0.14)	(0.39)		(−0.79)	(−0.02)	(−0.16)	(−0.49)	(1.68)	(0.27)	(0.34)	(0.94)
DEU	−0.05	0.10	**0.17**		**0.13**	0.01	0.11	**0.11**	0.17	0.07	0.15
	(−0.42)	(0.91)	(1.6)		(2.24)	(0.11)	(0.93)	(1.84)	(1.3)	(0.61)	(1.2)
ITA	−0.07	0.08	**0.18**	0.10		−0.02	−0.09	0.05	0.18	0.15	0.09
	(−0.77)	(0.71)	(1.73)	(0.77)		(−0.19)	(−0.61)	(0.62)	(1.28)	(1.29)	(0.8)
JPN	−0.01	0.09	**0.12**	−0.06	0.03		0.01	0.07	0.02	**0.12**	0.07
	(−0.1)	(0.91)	(1.62)	(−0.75)	(0.49)		(0.06)	(1.08)	(0.23)	(1.41)	(0.67)
NLD	0.09	**0.27**	**0.16**	0.06	0.06	0.03		**0.13**	**0.25**	0.17	**0.25**
	(0.86)	(2.48)	(2.15)	(0.68)	(1.18)	(0.52)		(2.21)	(2.17)	(1.35)	(2.2)
SWE	−0.05	**0.30**	0.11	0.15	**0.13**	0.02	0.05		0.15	**0.22**	**0.31**
	(−0.38)	(2.52)	(1.1)	(1.3)	(1.33)	(0.25)	(0.37)		(1.22)	(1.66)	(2.22)
CHE	0.02	0.05	0.02	−0.04	0.01	−0.04	−0.02	**0.13**		0.06	0.12
	(0.25)	(0.44)	(0.19)	(−0.37)	(0.27)	(−0.8)	(−0.22)	(2.64)		(0.63)	(0.98)
GBR	0.04	0.09	0.03	−0.09	0.00	0.05	−0.12	0.03	−0.05		0.10
	(0.53)	(0.84)	(0.33)	(−1.1)	(0.05)	(0.84)	(−1.12)	(0.67)	(−0.61)		(0.73)
USA	0.04	0.07	0.00	−0.05	0.03	−0.05	−0.04	0.05	−0.04	0.02	
	(0.43)	(0.61)	(−0.06)	(−0.75)	(0.79)	(−1.08)	(−0.48)	(1.18)	(−0.42)	(0.18)	
Average	0.00	0.11	0.08	−0.01	0.04	0.00	−0.03	0.08	0.07	0.08	0.12
Pooled	0.00	**0.11**	**0.09**	0.01	**0.05**	0.01	−0.01	**0.09**	0.06	0.09	**0.13**
	(−0.04)	(1.42)	(1.61)	(0.21)	(1.27)	(0.3)	(−0.19)	(2.38)	(0.78)	(1.33)	(1.75)

Table 5 continued

Country (i)	2000:9–2010:12										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		0.04	0.00	−0.01	−0.04	0.06	−0.01	0.11	**0.16**	0.02	0.09
		(0.31)	(0.03)	(−0.14)	(−0.49)	(0.7)	(−0.12)	(1.4)	(1.48)	(0.15)	(0.59)
CAN	0.18		**0.19**	**0.15**	0.10	**0.18**	0.12	**0.25**	**0.25**	**0.22**	**0.29**
	(1.29)		(1.74)	(1.76)	(0.95)	(1.85)	(1.2)	(3.59)	(2.12)	(1.56)	(2)
FRA	**0.37**	−0.01		0.14	−0.43	**0.17**	−0.21	0.18	**0.56**	0.15	0.31
	(1.56)	(−0.04)		(0.61)	(−2.05)	(1.38)	(−0.68)	(1.15)	(2.19)	(0.47)	(1.2)
DEU	**0.46**	0.07	0.03		−0.33	**0.23**	−0.25	0.20	**0.51**	0.21	**0.47**
	(1.86)	(0.4)	(0.11)		(−1.68)	(1.83)	(−0.92)	(1.06)	(1.92)	(0.72)	(1.81)
ITA	**0.57**	0.11	**0.69**	**0.34**		**0.20**	0.12	**0.28**	**0.54**	**0.46**	**0.50**
	(2.78)	(0.7)	(3.34)	(2.07)		(1.66)	(0.61)	(2.25)	(3.03)	(1.78)	(2.52)
JPN	0.06	0.11	0.08	0.10	0.01		0.09	0.04	**0.22**	0.08	0.06
	(0.4)	(0.77)	(0.75)	(1.19)	(0.07)		(0.83)	(0.39)	(1.7)	(0.56)	(0.46)
NLD	**0.43**	0.14	0.48	**0.46**	0.01	**0.28**		**0.34**	**0.68**	**0.52**	**0.59**
	(1.87)	(0.95)	(1.54)	(2.04)	(0.07)	(2.15)		(2.19)	(2.54)	(1.81)	(2.85)
SWE	0.23	0.11	0.08	0.05	−0.09	0.16	−0.08		0.24	0.04	**0.32**
	(0.92)	(0.59)	(0.37)	(0.25)	(−0.54)	(1.03)	(−0.4)		(1.23)	(0.15)	(1.42)
CHE	0.16	0.00	−0.11	−0.05	−0.18	**0.16**	−0.18	0.05		0.02	0.08
	(1.22)	(−0.05)	(−0.79)	(−0.38)	(−1.63)	(1.99)	(−1.22)	(0.49)		(0.13)	(0.58)
GBR	**0.27**	−0.02	0.12	0.16	−0.09	**0.14**	−0.10	**0.21**	**0.38**		**0.28**
	(1.63)	(−0.15)	(0.58)	(1.1)	(−0.55)	(1.48)	(−0.58)	(1.86)	(2.5)		(1.55)
USA	**0.30**	0.05	−0.04	−0.06	−0.16	**0.16**	−0.07	**0.19**	0.19	0.08	
	(1.58)	(0.24)	(−0.2)	(−0.44)	(−1.09)	(1.57)	(−0.38)	(1.65)	(1.33)	(0.31)	
Average	0.30	0.06	0.15	0.13	−0.12	0.17	−0.06	0.17	0.36	0.18	0.29
Pooled	**0.31**	0.08	**0.12**	**0.11**	−0.06	**0.19**	0.00	**0.18**	**0.33**	**0.17**	**0.26**
	(2.05)	(0.63)	(1.7)	(2)	(−0.75)	(1.95)	(−0.04)	(2.97)	(3.15)	(1.25)	(2.34)

Table 6 Pairwise Granger causality test for second subsample: 1982:9–2002:9 and 2002:10–2010:12

Country (i)	1982:9–2002:9										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		0.11	**0.15**	**0.18**	**0.08**	**0.11**	0.10	**0.09**	0.07	0.08	**0.13**
		(1.3)	(2.04)	(2.26)	(1.85)	(1.79)	(1.17)	(2)	(0.91)	(0.82)	(1.38)
CAN	−0.01		0.06	−0.01	0.02	0.02	−0.08	**0.09**	−0.04	0.01	0.03
	(−0.21)		(0.87)	(−0.11)	(0.34)	(0.38)	(−0.96)	(2.12)	(−0.39)	(0.13)	(0.34)
FRA	−0.01	0.02		−0.04	−0.01	0.01	−0.07	**0.14**	0.04	0.03	0.10
	(−0.12)	(0.21)		(−0.39)	(−0.15)	(0.17)	(−0.5)	(1.84)	(0.31)	(0.24)	(0.85)
DEU	−0.06	0.11	**0.21**		**0.12**	0.06	0.10	**0.15**	0.17	0.08	**0.16**
	(−0.51)	(1.15)	(1.99)		(2.09)	(0.8)	(0.83)	(2.27)	(1.36)	(0.7)	(1.48)
ITA	−0.08	0.06	**0.21**	0.12		0.00	−0.08	0.05	**0.17**	0.14	0.09
	(−0.88)	(0.6)	(1.99)	(1.04)		(0.01)	(−0.62)	(0.73)	(1.32)	(1.25)	(0.87)
JPN	−0.02	0.07	**0.12**	−0.04	0.03		0.02	0.07	0.04	**0.12**	0.05
	(−0.3)	(0.84)	(1.78)	(−0.5)	(0.55)		(0.32)	(1.17)	(0.52)	(1.47)	(0.52)
NLD	0.07	**0.26**	**0.22**	**0.15**	0.07	0.07		**0.17**	**0.27**	**0.19**	**0.28**
	(0.76)	(2.92)	(2.86)	(1.59)	(1.36)	(1.29)		(2.75)	(2.38)	(1.58)	(2.75)
SWE	−0.08	**0.26**	**0.14**	**0.17**	0.12	0.04	0.04		0.13	**0.21**	**0.29**
	(−0.6)	(2.36)	(1.5)	(1.56)	(1.27)	(0.49)	(0.32)		(1.13)	(1.63)	(2.16)
CHE	0.01	0.06	0.06	0.03	0.03	−0.01	0.01	**0.15**		0.08	0.14
	(0.14)	(0.64)	(0.68)	(0.29)	(0.65)	(−0.19)	(0.12)	(3.32)		(0.9)	(1.25)
GBR	0.03	0.08	0.06	−0.04	0.01	0.07	−0.10	0.05	−0.03		0.09
	(0.44)	(0.94)	(0.68)	(−0.57)	(0.24)	(1.26)	(−1)	(1.1)	(−0.41)		(0.76)
USA	0.03	0.10	0.03	−0.03	0.04	−0.02	−0.04	**0.08**	−0.03	0.02	
	(0.35)	(0.9)	(0.5)	(−0.4)	(0.94)	(−0.41)	(−0.45)	(1.67)	(−0.27)	(0.19)	
Average	−0.01	0.10	0.11	0.03	0.04	0.02	−0.02	0.10	0.07	0.09	0.12
Pooled	−0.01	**0.10**	**0.12**	0.05	**0.05**	0.04	0.00	**0.11**	0.07	**0.09**	**0.13**
	(−0.22)	(1.57)	(2.24)	(0.82)	(1.44)	(0.87)	(−0.05)	(3.16)	(1.02)	(1.46)	(1.93)

Table 6 continued

Country (i)	2002:10–2010:12										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		−0.07	−0.12	−0.09	−0.08	0.00	−0.03	0.08	0.12	−0.07	0.03
		(−0.37)	(−0.98)	(−0.96)	(−1)	(0)	(−0.27)	(0.91)	(0.99)	(−0.43)	(0.14)
CAN	**0.26**		**0.19**	**0.21**	0.13	0.09	**0.22**	**0.28**	**0.29**	**0.26**	**0.37**
	(1.63)		(1.64)	(2.42)	(1.07)	(0.81)	(2.07)	(3.88)	(2.48)	(1.82)	(2.48)
FRA	**0.48**	0.02		0.13	−0.32	0.03	0.06	0.21	**0.82**	0.31	**0.44**
	(1.89)	(0.1)		(0.56)	(−1.56)	(0.27)	(0.19)	(1.34)	(3.11)	(0.95)	(1.65)
DEU	0.49	−0.02	−0.34		−0.31	0.07	−0.14	0.06	**0.59**	0.21	**0.47**
	(1.74)	(−0.09)	(−1.02)		(−1.54)	(0.6)	(−0.48)	(0.27)	(1.93)	(0.74)	(1.62)
ITA	**0.70**	0.20	**0.72**	**0.27**		0.11	0.23	**0.34**	**0.69**	**0.56**	**0.65**
	(3.02)	(0.92)	(3.47)	(2.01)		(0.8)	(1.06)	(2.61)	(3.51)	(1.9)	(2.82)
JPN	0.18	0.22	0.06	0.09	0.00		0.11	0.05	**0.26**	0.08	0.15
	(1.04)	(1.12)	(0.51)	(0.96)	(0.01)		(0.82)	(0.47)	(1.84)	(0.48)	(0.92)
NLD	**0.40**	0.03	0.21	0.32	−0.07	0.13		0.23	**0.65**	0.39	**0.57**
	(1.53)	(0.13)	(0.57)	(1.35)	(−0.28)	(1.01)		(1.33)	(2.4)	(1.32)	(2.31)
SWE	0.28	0.10	0.09	0.09	−0.02	0.02	0.15		0.44	0.15	**0.37**
	(1.01)	(0.44)	(0.42)	(0.53)	(−0.11)	(0.11)	(0.81)		(1.91)	(0.62)	(1.59)
CHE	0.20	−0.03	−0.27	−0.20	−0.31	0.09	−0.16	−0.01		0.00	0.08
	(1.29)	(−0.19)	(−1.6)	(−1.41)	(−2.6)	(1.02)	(−0.98)	(−0.07)		(0)	(0.49)
GBR	**0.34**	−0.08	0.06	0.11	−0.10	0.04	−0.03	**0.22**	**0.44**		**0.37**
	(1.72)	(−0.44)	(0.3)	(0.8)	(−0.54)	(0.41)	(−0.16)	(1.9)	(2.67)		(1.71)
USA	0.29	−0.01	−0.10	0.02	−0.16	0.02	0.03	**0.21**	**0.26**	0.13	
	(1.33)	(−0.05)	(−0.59)	(0.16)	(−0.98)	(0.23)	(0.14)	(1.75)	(1.67)	(0.46)	
Average	0.36	0.04	0.06	0.10	−0.11	0.06	0.05	0.16	0.44	0.21	0.35
Pooled	**0.37**	0.07	0.07	**0.10**	−0.06	0.07	0.07	**0.18**	**0.41**	0.19	**0.33**
	(2.37)	(0.45)	(0.76)	(1.52)	(−0.63)	(0.76)	(0.67)	(2.74)	(3.38)	(1.3)	(2.53)

Table 7 Pairwise Granger causality test for second subsample: 1982:9–2007:5 and 2007:6–2010:12

Country (i)	1982:9–2007:5										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		0.10	**0.14**	**0.15**	**0.08**	**0.09**	0.09	**0.09**	0.08	0.09	**0.13**
		(1.25)	(2.03)	(2.34)	(1.87)	(1.71)	(1.27)	(2.08)	(1.04)	(1)	(1.44)
CAN	0.00		0.07	0.03	0.02	0.01	−0.02	**0.10**	0.01	0.04	0.08
	(−0.08)		(1.25)	(0.57)	(0.47)	(0.15)	(−0.26)	(2.51)	(0.08)	(0.55)	(0.91)
FRA	0.00	0.00		−0.03	−0.02	0.01	−0.06	**0.12**	0.07	0.04	0.11
	(0.04)	(0)		(−0.31)	(−0.37)	(0.11)	(−0.46)	(1.72)	(0.58)	(0.27)	(1.01)
DEU	−0.03	0.09	**0.18**		**0.10**	0.05	0.07	**0.13**	**0.21**	0.09	**0.19**
	(−0.26)	(0.92)	(1.74)		(1.68)	(0.81)	(0.65)	(2.02)	(1.66)	(0.75)	(1.63)
ITA	−0.05	0.06	**0.20**	0.13		0.01	−0.04	0.07	**0.18**	**0.15**	0.12
	(−0.63)	(0.66)	(2.12)	(1.23)		(0.07)	(−0.34)	(1.13)	(1.49)	(1.41)	(1.19)
JPN	−0.01	0.09	**0.13**	0.00	0.03		0.05	0.07	0.07	**0.13**	0.07
	(−0.21)	(1.14)	(1.99)	(−0.01)	(0.55)		(0.73)	(1.27)	(0.94)	(1.69)	(0.84)
NLD	0.08	**0.20**	**0.19**	**0.13**	0.05	0.06		**0.15**	**0.28**	**0.18**	**0.27**
	(0.83)	(2.24)	(2.51)	(1.49)	(0.91)	(1.25)		(2.5)	(2.61)	(1.52)	(2.72)
SWE	−0.06	**0.22**	0.10	0.11	0.10	0.03	0.01		0.11	0.17	**0.27**
	(−0.48)	(2.07)	(1.09)	(1.14)	(1.02)	(0.38)	(0.05)		(1)	(1.39)	(2.12)
CHE	0.02	0.03	0.04	0.02	0.02	0.00	0.00	**0.13**		0.08	0.13
	(0.21)	(0.35)	(0.46)	(0.21)	(0.32)	(−0.06)	(0)	(2.86)		(0.93)	(1.27)
GBR	0.04	0.05	0.05	−0.02	0.01	0.06	−0.09	0.05	−0.01		0.10
	(0.56)	(0.65)	(0.65)	(−0.27)	(0.12)	(1.16)	(−0.95)	(1.14)	(−0.12)		(0.93)
USA	0.03	0.04	0.03	−0.02	0.03	−0.03	−0.03	**0.07**	−0.01	0.03	
	(0.34)	(0.39)	(0.4)	(−0.35)	(0.76)	(−0.8)	(−0.47)	(1.66)	(−0.13)	(0.28)	
Average	0.00	0.08	0.10	0.04	0.03	0.02	−0.01	0.09	0.09	0.09	0.13
Pooled	0.00	**0.08**	**0.10**	0.05	0.04	0.03	0.00	**0.10**	**0.08**	**0.10**	**0.14**
	(−0.04)	(1.32)	(2.18)	(0.99)	(1.14)	(0.81)	(0.01)	(3.08)	(1.36)	(1.68)	(2.24)

Table 7 continued

Country (i)	2007:6–2010:12										
	$\hat{\beta}_{i,AUS}$	$\hat{\beta}_{i,CAN}$	$\hat{\beta}_{i,FRA}$	$\hat{\beta}_{i,DEU}$	$\hat{\beta}_{i,ITA}$	$\hat{\beta}_{i,JPN}$	$\hat{\beta}_{i,NLD}$	$\hat{\beta}_{i,SWE}$	$\hat{\beta}_{i,CHE}$	$\hat{\beta}_{i,GBR}$	$\hat{\beta}_{i,USA}$
AUS		−0.04	−0.26	−0.30	−0.23	0.05	−0.14	0.10	0.06	−0.26	−0.06
		(−0.88)	(−2.75)	(−1.28)	(0)	(0.36)	(−1.55)	(0)	(0.22)	(−1.35)	(−0.62)
CAN	0.19		0.05	0.13	0.03	**0.27**	0.04	**0.38**	0.18	0.08	0.25
	(0.78)		(0.46)	(0.62)	(0.35)	(2.46)	(0.22)	(3.63)	(0)	(0)	(1.59)
FRA	0.75	0.22		−0.03	−0.86	0.18	−0.09	0.28	1.30	0.39	**0.55**
	(3.08)	(0)		(0)	(−8.98)	(1.27)	(0)	(0)	(2.21)	(1.58)	(14.04)
DEU	0.68	0.14	−0.23		−0.39	0.23	−0.20	0.22	0.80	0.15	**0.47**
	(1.88)	(0)	(0)		(−3.21)	(1.44)	(−6.14)	(0.77)	(1.59)	(0)	(3.32)
ITA	**0.91**	0.40	**1.46**	**0.51**		0.32	0.14	0.40	1.39	0.79	0.84
	(3.91)	(0)	(12.27)	(13.24)		(2.19)	(0.51)	(0)	(0)	(2.42)	(0)
JPN	0.25	0.33	0.04	0.15	0.00		0.13	0.05	0.23	0.07	0.21
	(0.96)	(0)	(0.31)	(0.98)	(0.05)		(0.75)	(0)	(1.18)	(0)	(1.74)
NLD	0.68	**0.48**	**0.41**	**0.45**	0.11	**0.42**		0.47	**0.92**	0.57	**0.95**
	(2.66)	(3.94)	(3.45)	(3.3)	(0)	(3.31)		(0)	(3.8)	(2.78)	(11.28)
SWE	0.35	0.27	0.09	0.08	0.06	0.33	0.12		0.57	0.18	**0.39**
	(0.91)	(0)	(0.62)	(0.54)	(0.66)	(2.11)	(0.85)		(2.01)	(1.92)	(15.62)
CHE	0.20	0.05	−0.43	−0.33	−0.49	0.06	−0.26	0.06		−0.18	0.01
	(1.19)	(0.48)	(−6.27)	(−1.52)	(0)	(0.62)	(−1.14)	(0)		(−1.35)	(0.14)
GBR	0.48	0.02	−0.11	−0.08	−0.29	0.11	−0.19	0.23	**0.67**		**0.43**
	(1.66)	(0)	(−0.34)	(−0.91)	(−7.57)	(0.86)	(0)	(0)	(3.95)		(2.66)
USA	0.52	0.28	−0.21	−0.02	−0.26	0.21	0.01	0.35	0.44	0.02	
	(1.5)	(0)	(−7.84)	(−0.1)	(−2.38)	(1.66)	(0)	(1.9)	(1.29)	(0)	
Average	0.50	0.22	0.11	0.09	−0.21	0.21	−0.03	0.24	0.65	0.21	0.41
Pooled	**0.46**	0.23	0.02	0.05	−0.12	0.21	−0.02	**0.25**	**0.54**	0.14	**0.38**
	(2.53)	(1.19)	(0.14)	(0.43)	(−1.11)	(1.46)	(−0.1)	(2.43)	(2.38)	(0.77)	(2.46)

References

Ang A, Bekaert G (2007) Stock return predictability: is it there? Rev Financ Stud 20:651–707
Armah NA, Swanson NR (2015) Predictive inference under model misspecification. In: Rapach DA, Wohar ME (eds) Forecasting in the presence of structural breaks and model uncertainty. pp 195–230, published online 10 Mar 2015
Aye GC, Deale FW, Gupta R (forthcoming) Does debt ceiling and government shutdown help in forecasting the US equity risk premium? Panoeconomicus
Bai J, Perron P (2003) Computation and analysis of multiple structural change models. J Appl Econom 18(1):1–22
Balcilar M, Ozdemir ZA (2013) The export-output growth nexus in Japan: a bootstrap rolling window approach. Empir Econ 44:639–660
Balcilar M, Ozdemir ZA, Arslanturk Y (2010) Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. Energy Econ 32:1398–1410
Barnett A, Mumtaz H, Theodoridis K (2012) Forecasting UK GDP growth, inflation and interest rates under structural change: a comparison of models with time-varying parameters. Bank of England Working Papers 450, Bank of England
Bossaerts P, Hillion P (1999) Implementing statistical criteria to select return forecasting models: what do we learn? Rev Financ Stud 12:405–428
Campbell JY (1987) Stock returns and the term structure. J Financ Econ 18:373–399
Campbell J, Thompson SB (2008) Predicting the equity return premium out of sample: can anything beat the historical average? Rev Financ Stud 21:1509–1531
Campbell JY, Shiller R (1998) Valuation ratios and the long-run stock market outlook. J Portf Manag 24(2):11–26
Clark TE, McCracken MW (2004) Evaluating long-horizon forecasts, Working Paper, University of Missouri-Columbia. ftp://ftp.soc.uoc.gr/students/aslanidis/Nektarios/Andrea/Clark%20&%20McCracken%20(2004).pdf
Dangl T, Halling M (2012) Predictive regressions with time-varying coefficients. J Financ Econ 106:157–181
Fama EF, Bliss RR (1987) The information in long-maturity forward rates. Am Econ Rev 77:680–692
Fama EF, French KR (1989) Business conditions and expected returns on stocks and bonds. J Financ Econ 25(1):23–49
Ferreira MA, Santa-Clara P (2011) Forecasting stock market returns: the sum of the parts is more than the whole. J Financ Econ 100:514–537
Ghosh S (2011) On the grouped selection and model complexity of the adaptive elastic net. Statistics and Computing 21:451–462
Goyal A, Santa-Clara P (2003) Idiosyncratic risk matters! J Finance 58:975–1007
Goyal A, Welch I (2008) A comprehensive look at the empirical performance of equity premium prediction. Rev Financ Stud 21:1455–1508
Granger CWJ (1996) Can we improve the perceived quality of economic forecasts. J Appl Econom 11(5):455–473
Groenewold N, Fraser P (1999) Time-varying estimates of CAPM betas. Mathematics and Computers in Simulation 48(4–6):531–539
Gupta R, Modise MP (2012) South African stock return predictability in the context data mining: the role of financial variables and international stock returns. Econ Model 29:908–916
Henkel SJ, Martin JS, Nadari F (2011) Time-varying short-horizon predictability. J Financ Econ 99:560–580
Hong H, Torous W, Valkanov R (2007) Do industries lead stock markets? J Financ Econ 83:367–396
Inglesi-Lotz R, Balcilar M, Gupta R (2014) Time-varying causality between research output and economic growth in US. Scientometrics 100:203–216
Kandel S, Stambaugh RF (1996) On the predictability of stock returns: an asset allocation perspective. Journal of Finance 51:385–424
Keim DB, Stambaugh RF (1986) Predicting returns in the stock and bond markets. J Financ Econ 17(2):357–390
Pesaran M, Timmermann A (1999) Model instability and choice of observation window, University of California at San Diego, Economics working paper Series, Department of Economics, UC San Diego

Pesaran MH, Timmermann A (2005) Small sample properties of forecasts from autoregressive models under structural breaks. Journal of Econometrics 129:183–217

Rapach DE, Wohar ME (2006) In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. J Empir Finance 13:231–247

Rapach DE, Zhou G (2013) Forecasting stock returns. In: Elliott Graham, Timmermann Allan (eds) Handbook of Economic Forecasting, vol 2. Elsevier, Amsterdam

Rapach DE, Strauss JK, Zhou G (2013) International stock return predictability: What is the role of the United States? J Finance LXVIII(4):1633–1662

Rizova S (2010) Predictable trade flows and returns of trade-linked countries. Working paper, University of Chicago, UT Dallas

Rozeff M (1984) Dividend yields are equity risk premiums. J Portf Manag 11:68–75

Stock JA, Watson MW (1996) Evidence on structural stability in macroeconomic time series relations. J Bus Econ Stat 14:11–30

Stock JH, Watson MW (2003) Forecasting output and inflation: the role of asset prices. J Econ Lit 41:788–829

Toda HY, Yamamoto T (1995) Statistical inference in vector autoregressions with possibly integrated processes. J Econom 66:225–250

Zou H, Zhang HH (2009) On the adaptive elastic-net with a diverging number of parameters. Ann Stat 37:1733–1751