Implications of Cointegration for Forecasting: A Review and an Empirical Analysis

Seyed Mahdi Barakchian*

Abstract

Cointegration has different theoretical implications for forecasting. Several empirical studies have compared the out of sample forecasting performance of cointegrted VECMs against unrestricted VARs in levels and in differences. The results of these studies have been generally mixed and inconclusive. This paper provides a comprehensive review over the subject, and also examines the effects of cointegration rank restrictions on forecasting performance of VAR models through conducting an empirical exercise in the framework of a new two-country (Canada-US) model. The results show that a VAR/DVAR model forecasts as well as the best cointegrated VAR model (and even better). Therefore, it seems that using cointegration techniques does not pay a dividend.

Key Words: *Cointegration, Forecasting using VECX*, Rank restrictions*

JEL Classifications: *C32, C51, C53*

---

* Graduate School of Management and Economics, Sharif University of Technology, barakchian@sharif.edu.

1. Introduction

The fall of Keynesian macroeconomic theory and poor forecasting performance of structural models based on Keynesian system of equations paradigm in 1970s caused a radical change of direction and shifted macroeconomic forecasting towards nonstructural forecasting. Following that, due to some important advances in econometric theory, particularly the Rational Expectations revolution (Lucas, 1976), and the introduction of Vector Autoregressive (VAR) models (Sims, 1980) and cointegration analysis (Engle and Granger, 1987, and Johansen, 1988, 1991), essential changes have occurred in macroeconometric modeling and forecasting. For a review on the history of macroeconomic forecasting, see Diebold (1998).

The theoretical implications of cointegration for forecasting can be grouped into three competing conjectures, based on my understanding of the literature.

First Conjecture: It is generally believed that cointegrating restrictions, when they really hold, improve forecast accuracy in the medium to long run. This belief stems from two theoretical results. First, long run forecasts generated from cointegrated systems would satisfy the cointegrating relations exactly. Second, the cointegrating relations have finite error variance for their long run forecasts. The first point will be formally illustrated in Section 2 and for the second point see, for example, Christoffersen and Diebold (1998).

Second Conjecture: Information about whether the cointegrating relations are in equilibrium or not and if not, how far they are from the equilibrium, helps to predict where the variables move to in the near future since the deviations from the equilibrium tend to be eliminated. Therefore, the values of error correction terms are valuable for the near-horizon forecasting of the system. But, since the long run forecasts of the error correction terms are always zero, it is doubtful that the error correction terms can provide any information to improve long run forecasting; see Christoffersen and Diebold (1998).

Third Conjecture: The third conjecture has been provided by Clements and Hendry, who have developed a framework for economic forecasting in a series of papers and books (for example, Clements and Hendry, 1996 and 1999, and Hendry, 2006 inter alia).[1] They argue that the theory of economic forecasting, or so-called 'textbook theory', historically has relied on two key assumptions:

1. The model is a good representation of the economy; and
2. The structure of the economy will remain relatively unchanged (Hendry and Clements, 2003), but empirical evidence suggests that these two assumptions need not hold in practice. Instead, Clements and Hendry propose two matching assumptions:

- Models are simplified representations which are incorrect in many ways; and
- Economies both evolve and suddenly shift (Hendry and Clements, 2003).

They argue that the implications of this 'more realistic setting' are almost contrary to the implications of the textbook theory and they match the empirical evidence. They show that among the nine potential sources of forecast failure, shifts in deterministic terms, i.e. shifts in intercepts and trend coefficients, are the most pernicious ones for forecasting,[2] and these shifts are frequent in economics.

The new framework has also some implications for forecasting performance of cointegrated Vector Error Correction Models (VECMs). Clements and Hendry (1999) show that equilibrium-mean shifts are one sort of deterministic shifts, and therefore, could induce major forecast failures. Thus, imposing cointegration improves forecasting accuracy of the model so long as there is no shift in equilibrium means. But if a structural break occurs and the equilibrium mean shifts, say, upwards then the VECM interprets this shift as a disequilibrium. Since the VECM is designed to remove any disequilibrium by adjusting in the opposite direction in order to put the time series back on track,

1. Hendry and Clements (2003) provide an excellent survey. The review of Clements and Hendry's works that we draw here is partly based on Hendry and Clements' (2003) survey.
2. Clements and Hendry (1998, 1999) develop a forecast-error taxonomy by which they distinguish between nine sources of forecast errors.

the VECM forecasts a fall in the data when the data is actually jumping up, and conversely.[1]

Several studies have been carried out in recent years which compare the out of sample forecasting performance of cointegrated VECMs against unrestricted VARs in levels and in first differences (see Section 2 for the review). The results of these studies have been generally mixed and inconclusive. This is partly due to the fact that forecast outcomes are case- and period-specific. This can also be partly attributed to the uncertainties that surround the estimation of a cointegrated VECM. It is widely known that cointegration rank estimates are very sensitive to the number of lags and to the sample size. Moreover, when the characteristic roots of the process are close to one, the cointegration rank tests are not powerful enough to reject the unit root hypothesis. These, therefore, may lead to mis-detecting the correct number of unit roots in the VAR, which in turn could possibly result in poor forecasting.[2]

The fact that cointegrated VECMs do not make particularly good forecasting models can also be accounted to some extent by error in estimation of the parameters of cointegrating vectors. If that is the case, specifying cointegrating relations based on economic theory rather than estimating them may provide an improvement in forecast accuracy.[3] Few papers have examined forecasting performance of over-identified cointegrated VECMs in which the long run relations motivated by economic theory are imposed on the cointegrating relations of the model; see, for example, Hoffman and Rasche (1996), Anderson et al. (2002) and Lastrapes (2002). However, since these papers typically compare between forecasting performance of ‘over-identified’ cointegrated VECMs (which impose both rank restrictions and long run theory restrictions) and forecasting performance of unrestricted VARs (which impose neither rank restrictions nor long run theory restrictions), it is impossible to separate the contributions of cointegration rank restrictions from the effects of

1. Clements and Hendry (1996) also show that VARs in differences forecast better than cointegrated-based VECMs when the means of the cointegrating relations are non-constant.
2. There are also other sources of uncertainty. For example, if one is interested in forecasting output growth then he has to decide how many other variables should be included in the VAR which clearly has not a single answer.
3. Alternatively, using Bayesian type parameter restrictions could be worth pursuing.

imposing over-identifying restrictions provided by economic theory. For more details on the role of long run theory restrictions in forecasting, see Barakchian (2012a).

In this paper we follow a twofold object. First, and the major, one is to provide a comprehensive review over the subject of ‘cointegration and forecasting’. Second one is to examine the effects of cointegration rank restrictions on forecasting performance of VAR models through conducting an empirical exercise.

We conduct this exercise in the context of a two-country model consisting of cointegrated VECMs of the US and Canadian economies. We call this two-country model a VECX*(r,r*), where r and r* refer to the cointegration ranks of the models of Canada and the US, respectively. Our focus being on Canada, we treat the US as ‘the rest of the world’ for Canada in this two-country model.

We believe the two-country model is well suited to study the effects of cointegration rank restrictions on forecasting performance of VECMs. Firstly, it allows for a large number of cointegrating relations. In this exercise, we consider up to five cointegrating relations in the model of Canada and up to two cointegrating relations in the model of the US which sums up to seven cointegrating relations in the full system, i.e. in the VECX*(r,r*). Secondly, there is a long and reliable time series available for US and Canadian macroeconomic variables which are required for such a study.

The results show that as the cointegration ranks (r,r*) imposed on the models are increased, the accuracy of the forecasts of output, short and long term interest rates typically decrease. But this pattern does not hold for inflation. We also find that a VAR model for forecasting inflation and a DVAR[1] model for forecasting the rest of the variables work as well as the best VECX* model (and even better), and therefore, it seems that using cointegration techniques, which is not simple/easy, does not pay a dividend.

The outline of the paper is as follows. Next section reviewes the papers which examine the effects of cointegration on forecasting performance of VAR

1. A ‘Var in Differences’ is called a DVAR.

models. Section 3 describes the two-country model. Section 4 examines the forecasting performance of the exactly-identified VECX*(r,r*) models. The last section concludes.

2. Cointegration and Forecasting: A Review of Empirical Studies

Consider this VECM representation of a VAR model:[1]

$$\Delta z_t = \Pi\left(z_{t-1} + \gamma(t-1)\right) + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + a_0 + u_t, \quad (1)$$

where z_t is an m×1 vector of I(1) variables, a_0 is an m×1 vector of deterministics, Γ_i is an m×m matrix of parameters, γ is an m×1 vector of parameters, and $u_t \sim iid(0, \Sigma)$, where Σ is a positive definite matrix. If the $rank(\Pi) = r$, where $0 < r < m$ then the matrix Π can be decomposed as $\Pi = \alpha\beta'$ where α and β are m×r full column rank matrices. We can rewrite the equation as:

$$\Delta z_t = \Pi_* z_{t-1}^* + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + a_0 + u_t, \quad (2)$$

where $\Pi_* = \alpha\beta_*', \beta_*' = [\beta' \quad \beta'\gamma]$, and $z_{t-1}^* = (z_{t-1}, t-1)'$. Each column of β_* represents one cointegrating vector and α is the matrix of loading coefficients.

To see how cointegration will affect long run forecasting, we define $z_{T+h|T}$ as the h-step ahead forecast formed at the forecast origin T. Using the VECM representation, the h-step ahead forecast of Δz_t will be

$$\Delta z_{T+h|T} = \Pi_* z_{t+h-1|T}^* + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t+h-i|T} + a_0 + u_{T+h|T}. \quad (3)$$

u_t is white noise, therefore $u_{T+h|T} = 0, \forall h > 0$. In addition, since $\Delta z_{T+h|T}$ is stationary, then $\lim_{h\to\infty} \Delta z_{T+h|T} = \mu$ where $\mu = E(\Delta z_t)$. Therefore as $h \to \infty$ the above forecast relation can be written as:

$$\mu = \lim_{h\to\infty} \Pi_* z_{t+h-1|T}^* + \sum_{i=1}^{p-1} \Gamma_i \mu + a_0. \quad (4)$$

1. In this example, intercepts are unrestricted but the trend coefficients are restricted to lie in the cointegrating space (following Johansen's (1996) case IV) in order to prevent having quadratic trends in the level of the variables.

This implies:

$$\lim_{h\to\infty} \beta_* z^*_{t+h-1|T} = C \qquad (5)$$

where $C = (\alpha'\alpha)^{-1}\alpha'\left[\left(I - \sum_{i=1}^{p-1} \Gamma_i\right)\mu - a_0\right]$ is a constant vector. As it is obvious from the above equation, each cointegrating relation implies a linear constraint on the long horizon forecasts of z_t. Therefore, it is expected that using cointegrating relations would improve forecast accuracy of the long horizon forecasts in a cointegrated system. This is a formal demonstration of the first point of the first conjecture noted in Introduction. Clearly this holds if the model and the restrictions are correct and there are no major structural breaks.[1]

We start the review with Clements and Hendry (1995), one of the major early studies in this area. In a Monte Carlo experiment, Clements and Hendry (1995) show that:

1. An estimated cointegrated model outperforms a DVAR in forecasting the level of the variables, the first difference of the variables and the cointegrating relations, at all forecast horizons.[2] However, as forecast horizon increases, their difference becomes negligible (except for the forecasts of the cointegrating relation).[3]
2. There is little difference between the cointegrated model and unrestricted VAR (UVAR) in forecasting at short horizons (up to *5*-steps ahead). But at long horizons, there is a clear advantage to use the cointegrated model. This is accentuated for the smaller samples.[4] This is consistent with the

1. According to Lin and Tsay (1996), cointegration improves forecasting in two ways: first in improving parameter estimation, and second in improving long term forecasts.
2. They consider up to *20*-steps ahead horizon.
3. Throughout the review, we use the terms of "short", "medium", and "long" forecast horizons exactly as they have been used in the papers, so it may seem that these terms do not refer to the same horizons across different papers. However, in each case we indicate that the term refers specifically to how many steps ahead forecasts.
4. The smaller sample includes 50 observations while the larger one includes 100 observations.

findings of Engle and Yoo (1987). But there is very little to choose between the cointegrated model and UVAR when forecasting Δx_t and the cointegrating relations.

3. In terms of Generalized Forecast Error Second Moment (GFESM),[1] there is little to choose between the cointegrated model and UVAR, while DVAR is apparently inferior.
4. In an empirical study using UK money data, they show that in terms of Root Mean Square Forecast Error (RMSFE),
5. UVAR does almost as well as the cointegrated model in forecasting $x_t, \Delta x_t$ and the cointegrating relation at all forecast horizons (it just fares a little worse).
6. DVAR predicts worse than the cointegrated model. However, the losses from using DVAR in forecasting Δx_t almost disappear at longer forecast horizons (*12*-quarters ahead). One anomaly between the Monte Carlo and the empirical results is that the DVAR does as well as the cointegrated model in forecasting the cointegrating relation.

Therefore, they conclude that the benefit from imposing reduced-rank cointegration restrictions is little compared to unrestricted VAR unless the sample size is small. They add that in larger systems, however, imposing cointegration restrictions may improve forecasting performance. They also conclude that 'imposing too few cointegration vectors may impose greater costs in forecast accuracy than allowing the presence of 'spurious' levels terms.'

In another major study, Lin and Tsay (1996) in a set of Monte Carlo experiments use Johansen procedure to determine the cointegration rank and estimate the parameters of the model. Their Monte Carlo simulation shows that:

1. Almost always, imposing the correct number of unit roots on the system yields the best forecasting performance.[2]
2. Forecast accuracy declines as the specified number of unit roots moves away from the correct one. Under-estimating the number of unit roots

1. GFESM has been proposed by Clements and Hendry (1993).
2. They consider up to *60*-steps ahead horizon.

may give poor long term forecasts (after almost *10*-steps ahead horizon).

3. If the Data Generating Process (DGP) is non-stationary, then over-estimating the number of unit roots might not reduce accuracy for long term forecasting. But in the case of a stationary time series, over-estimating the number of unit roots gives poor long term forecasts.
4. In a large sample (with 400 observations), when the roots of the process are on the unit circle or close to one, the specification of the number of unit roots is not critical for short term forecasts (such as *2*-steps ahead forecasts).

Using a range of macroeconomic and financial data sets, Lin and Tsay (1996) find that unit root constraints yield poor forecasts in some cases.[1] But for most cases, for long term forecasts (more than *15*-months ahead) unit root constraints can substantially improve forecast accuracy, whereas their effects for short and medium term forecasts (up to *10*-months ahead) are moderate.[2] They also find that the number of unit roots detected by the cointegration rank tests does not necessarily give the best forecasts. They conclude that:

> '*some minor deviations in the number of unit roots would not appreciably affect the forecasts when the series contain unit roots. Consequently, for practical forecasting, it is important to determine the stationarity of a time series, but is not critical to identify the exact number of unit roots in the data for nonstationary series*.'

While the work of Lin and Tsay (1996) was mostly focussed on the effect of specifying the number of unit roots in a system on forecasting performance, Christoffersen and Diebold (1998) concentrate on isolating the effects of cointegration from integration on forecasting performance in the following sense: They compare two forecasts of a multivariate cointegrated system; first,

1. Lin and Tsay (1996) use the term "unit root constraints", i.e. cointegration rank constraints, for the case where the number of unit roots detected by the cointegration tests are imposed on the VAR model.
2. It means that the difference in "the trace of the error covariance matrix" of different models is less than 10%.

forecasts from the multivariate model and, second, forecasts from the implied univariate representations of the variables where their integration properties are correctly specified.[1] They show analytically and also with a Monte Carlo simulation (using the trace MSFE ratio) that at long horizons the univariate forecasts perform as accurate as the cointegrated system forecasts; and imposition of cointegrating relations improves short term forecasts, so long as the number of unit roots in the system is identified correctly. They also show that univariate Box-Jenkins forecasts hang together, i.e. they preserve the cointegration property, if the variables are in fact cointegrated.[2] Therefore they conclude that 'there is no long-horizon benefit from imposing cointegration; all that matters is getting the level of integration right'. This result contradicts the finding of some other studies, e.g. Engle and Yoo (1987), who show that imposing cointegrating constraints tend to improve long horizon forecast accuracy. Christoffersen and Diebold (1998) argue that their set up has allowed them to isolate the effects of cointegration from integration while in others, the forecasts from a cointegrated system (which imposes both integration and cointegration) are compared with the forecasts from a VAR in levels (which imposes neither integration nor cointegration). They conclude that the forecasting performance of the VAR in levels is poor not because it does not impose cointegration but because it fails to impose integration.[3]

In another Monte Carlo experiment based on a fully specified economic model, Söderlind and Vredin (1996) examine the effects of cointegration

---

1. The univariate representation is modeled as an ARIMA (0, 1, n) using Wold's decomposition theorem. They emphasize that both models are correctly specified 'from a *univariate* perspective', but in the first one, the cointegrating restrictions are imposed and correlation between error terms across equations is allowed, while this is not the case in the second one.
2. They conclude that:

   '*it is in fact not imposition of cointegration on the forecasting system that yields the finite variance of the cointegrating combinations of the errors; rather it is the cointegration property inherent in the system itself, which is partly inherited by the correctly specified univariate forecasts*.'
3. To explain this note more, they compare the forecasts from a VAR in differences (DVAR) to the forecasts from the cointegrated system and show that at long horizons (using trace MSFE) the DVAR, 'which imposes integration but completely ignores cointegrtaion' forecasts as accurate as the cointegrated system.

restrictions on forecasting when the true cointegrating relations are imposed on the model. They find that the cointegrated VECM generates somewhat better forecasts than the unrestricted VAR, although the differences are small at short forecast horizon (*1*-step ahead) and also for the forecasts of the growth rates of the variables.[1] But the cointegrated VECM is outperformed by random walk model in some cases, especially in the case of growth rates. However, the performance of the VECM improves as the sample becomes longer. Söderlind and Vredin (1996) eventually conclude that 'cointegration analysis add little to forecasting'.

In an attempt to study the effects of cointegration in the context of a more volatile, high-frequency data set of exchange rates, McCrae et al. (2002) use daily data of Asian exchange rates and find that neither the ARIMA models nor the cointegrated based VECMs generate particularly accurate forecasts. In terms of RMSFE and mean absolute forecast error (MAFE), while generally the cointegrated based VECM does not perform significantly better than the ARIMA model for short term forecasts (up to five days), cointegration improves forecast accuracy over the medium forecast horizon (six to forty days). In terms of trace MSFE, the VECM outperforms ARIMA over both short and medium forecast horizons.

Wang and Bessler (2004) employ cointegration technique in a quite different context, i.e. agriculture, using a long time series of the annual US hog data. They compare forecasting performance of a range of unrestricted VARs and also reduced rank models. They find that in terms of the total of the RMSFE of the individual series,[2] the cointegrated based VECM outperforms all other models,[3] whereas the DVAR model performs poorly. Thus they conclude that 'it is more important to correctly specify the cointegration rank than to difference the possibly nonstationary data.' When they use the Clements and

1. They consider up to 8-steps ahead horizon.
2. This measure simply is equal to the sum of RMSFE of individual variables.
3. They note that since the orders of magnitude of the individual series are different, the results based on the total RMSFE as the overall measure of forecasting performance should be interpreted cautiously.

Hendry's (1998) encompassing test, the picture of the results becomes slightly different. The VECM does not encompass any model and no other model encompasses the VECM either. But in contrast, the VAR(2) model encompasses several other models at short forecast horizons (up to 2-years ahead horizons).

2.1. The Role of Long Run Theory Restrictions for Forecasting

Among many studies that examine the effects of cointegration modelling on forecasting, few of them have paid attention to the role of economic theory in specifying the cointegrating relations. Swanson (2002) provides a theoretic basis why imposing long run relations suggested by economic theory on cointegrating vectors could possibly improve forecasting performance of a cointegrated VECM. He employes Monte Carlo experiments in order to examine the conjecture that parsimonious models often perform better in terms of forecasting than more heavily parametrized but correctly specified alternative models, due to parameter estimation error. In one Monte Carlo study, Swanson (2002) shows that when DGP is a cointegrated VECM, for small samples, forecasts from a VAR in differences tend to be as well as those from a cointegrated VECM where the cointegration rank and the parameters of the cointegrating vectors are estimated (i.e. the true ones are not imposed on the model). However, as the sample size increases, the VECM forecasts dominate the VAR in differences. He attributes this result to errors in estimating cointegration rank and cointegrating vectors. He concludes that the evidence that cointegrated VECMs do not make particularly good forecasting models can be partly attributed to parameter estimation error. Consequently, specifying cointegrating relations based on economic theory rather than estimating them may provide an improvement in forecast accuracy.[1]

Hoffman and Rasche (1996) is one of the few studies in which over-identifying restrictions suggested by economic theory are imposed on the cointegrating relations of the VECM. They use US macro data to assess the forecasting performance of a cointegrated VECM against VARs in levels and in first differences. Three cointegrating relations of the VECM correspond to the

1. Of course, this can be true if the theory restrictions are valid.

money demand, the Fisher equation, and the interest rate differential.[1] Using GFESM, they find that there is some advantage to incorporating cointegration especially as compared to DVAR. When using RMSFE, the result is somewhat different. First, the pattern of the forecasting performance of the different models across the variables is quite diverse. Second, the advantage to incorporating cointegration, if any, is not confined to any particular representation of the data; i.e. it is not confined to level representation, to differences, or to cointegrating combinations.[2] Third, if incorporating cointegration is advantageous, it accrues only at longer forecast horizons (typically more than *12*-quarters ahead). Fourth, imposing the Fisher equation and interest rate differential on the model improves forecasts of inflation rate, the real interest rate and the risk premium. In contrast, imposing the money demand relation on the model does not improve forecasts of the money demand cointegration relation.

Anderson et al. (2002) use US macro data to estimate a VECM with four over-identified cointegrating relations corresponding to a long run money demand, a term structure and two Fisher equations. They show that when judged by GFESM, the cointegrated VECM delivers superior performance compared to a naive random walk model especially at short forecast horizons (particularly up to 4-quarters ahead)[3]. But when judged by RMSFE, the random walk model produces forecasts of comparable average accuracy. They also find that the forecasting performance of the cointegrated VECM is comparable to government agency and private sector forecasts (including Greenbook forecasts). They conclude that '*imposing some long-run structural (or balanced-growth) relationships may pay dividends*.'

---

1. The interest rate differential relation is defined as the difference between risky and risk-free short term interest rates, i.e. the difference between commercial paper rate and Treasury bill rate.
2. This is in contrast to Clements and Hendry (1995) who found that the relative advantage of cointegrtion accrues in level representations of data, but not in differences and cointegrating combinations representations.
3. They generate forecasts up to 16-quarters ahead horizon. As forecast horizon increases, the superiority of the cointegrated VECM decreases, though not monotonically.

In his comments on Anderson et al. (2002), Lastrapes (2002) questions the choice of GFESM as a criteria to compare forecast accuracy. According to him, while GFESM places the same weight on all variables in the system, perhaps the forecasts of some variables like output growth and inflation are much more important for the public/private forecasters. Frankly the performance of the cointegrated VECM is even worse than the alternatives (in terms of RMSFE) when forecasting the individual series of output and inflation in Anderson et al. (2002) study. Lastrapes uses encompassing tests to see if the over-identified cointegrated VECM estimated in Anderson et al. (2002) contains additional information over its competitors to improve forecast accuracy. He finds that although the cointegrated VECM is not superior to the random walk model in forecasting output in terms of RMSFE, but it could have additional information to forecast output which is not contained in the random walk model; simply it is not encompassed by the random walk model.

Jacobson et al. (2001) use a data set on the Swedish economy and show that the performance of the cointegrated VECM is improved in forecasting inflation, in terms of RMSFE criteria, if the long run theory restrictions are imposed on the cointegrating relations rather than relying on exactly identified cointegrating relations. The first long run relation considered is PPP. For the second and third long run relations they assume that the domestic and foreign interest rates are stationary.

Finally, in two related papers, Assenmacher-Wesche and Pesaran (2008) and Pesaran et al. (2009) show that pooled forecasts generated from a range of exactly-identified and over-identified VECMs estimated for the Swiss and global economies, respectively, beat naive forecasting models [Random Walk and AR(1)].

As it can be seen from the review, the papers have not arrived at the same conclusion when examining the effects of cointegration modeling on forecasting performance of VAR models; some have judged in favor of cointegration and some against. In the rest of the paper, we try to shed some light on the issue by conducting an empirical study in the context of a two-country model which comprises US and Canada.

3. Outline of the Two-country Model

In the two-country model, the US is modeled as a closed economy,[1] and includes five endogenous variables, $\boldsymbol{x}_t^*$: US short term interest rate, US long term interest rate, US inflation, US income per capita, and oil price. Canada is modeled as a small open economy and includes five endogenous variables, $\boldsymbol{x}_t$: Canadian real effective exchange rate, Canadian short term interest rate, Canadian long term interest rate, Canadian inflation, and Canadian per capita income. It also incldes $\boldsymbol{x}_t^*$ as a vector of weakly exogenous variables.[2] Long run interest rate is added to allow for yield curve relationship since macroeconomic variables and yield curve are mutually interdependent. Not only macro variables affect yield curve (Diebold et al., 2006), but yield curve can help explain changes in macro variables like GDP (Estrella and Mishkin, 1998). Pesaran et al. (2009) show that including long term interest rate contributes substantially to the forecast accuracy of output growth for Canada and the US.

We treat the US as the rest of the world for Canada in the two-country model and we justify it based on the fact that the US is the major partner of Canada with more than 70% share in trade of Canada.[3] Therefore, the US plays the role of a dominant economy for Canada and the effects of other countries on Canada are negligible compared to the US. For more details on the role of US variables in explaining variation of Canadian variables, see Barakchian (2012b).

To construct the two-country model, first two separate cointegrated VECMs are estimated for Canada and the US. These are quarterly models which are estimated over 1958Q1-1990Q4. Then these two models are combined to form a two-country model. We call the resultant model a VECX*(r,r*), where r and r* refer to the cointegration ranks of the models of Canada and the US, respectively.

1. In modeling the US as a closed economy we follow the sizable literature on the US economy. See, for example, Anderson et al. (2002), Bernanke and Mihov (1998), Christiano et al. (1996,1999) and Sims and Zha (2006) inter alia.
2. More details on the definitions and sources of the data are provided in the Appendix.
3. Over the period 2004-2008, the share has been 73% (on average). For more details, see 'Statistics Canada' at http://www40.statcan.gc.ca/l01/cst01/gblec02a-eng.htm.

Performing ADF and Phillips-Perron unit root tests suggest that the null hypothesis that all the ten variables contain a unit root can not be rejected, and thus the variables can be treated as I(1) variables. Based on the Akaike criterion, two lags are selected for the underlying unrestricted VARs of Canada and the US.[1]

3.1. Model for Canada

We estimate a cointegrated VARX* model for Canada, where both the endogenous, $\boldsymbol{x}_t$, and exogenous, $\boldsymbol{x}_t^*$, variables are included with two lags, over the period 1958Q1-1990Q4. In order to prevent having quadratic trends in the level of the variables in both models of Canada and the US, deterministics are treated according to the case IV, as described in Pesaran et al. (2000), in which, intercepts are unrestricted, but the trend coefficients are restricted to lie in the cointegrating space. The model for Canada can be written as

$$\Delta x_t = \Pi\big(z_{t-1} + \gamma(t-1)\big) + \Lambda\Delta x_t^* + \Psi\Delta z_{t-1} + a_0 + u_t, \qquad (6)$$

where x_t is a m×1 vector, x_t^* is a m*×1 vector and z_t is a $\widetilde{m} \times 1$ vector, where $z_t' = (x_t', x_t^{*\prime})$ and $\widetilde{m} = m + m^*$. The foreign variables, x_t^*, are treated as weakly exogenous.

When the rank of the matrix Π is $r < m$, Π can be decomposed as $\Pi = \alpha\beta'$, where α is a $m \times r$ matrix of loading coefficients and β is a $\widetilde{m} \times r$ matrix of cointegrating vectors. Even if we know an estimate of Π, without having more restrictions, α and β are not identifiable. Any non-singular $r \times r$ matrix Q can be used to write:

$$\Pi = \alpha\beta' = \big(\alpha Q'^{-1}\big)(Q'\beta') = \tilde{\alpha}\tilde{\beta}'$$

where $\tilde{\alpha} = \alpha Q'^{-1}$ and $\tilde{\beta} = \beta Q$ are observationally equivalent to α and β. So in order to identify α and β we need to impose r restrictions on each cointegrating vector, r^2 restrictions altogether. This will result in an exactly-identified set of cointegrating vectors. Johansen (1996) introduces a procedure for estimating α and β in which, the maximum likelihood estimate of β is computed as the

1. The results for the unit root tests and also for the lag order selection criteria are available upon request.

r eigenvectors corresponding to the first r largest eigenvalues of the canonical correlation matrix. These restrictions, which sometimes are called 'statistical', provide the r^2 restrictions needed for exact identification of β. In this exercise, we use the Johansen method to provide exact identifying restrictions. However, under exact identification, the forecasts are invariant to any non-singular transformation of the cointegrating relations.

3.2. Model for the US

Similarly, we estimate a cointegrated VAR for the US, where the variables are included with two lags, over 1958Q1-1990Q4. The US model can be written as

$$\Delta x_t^* = \Pi^*\left(x_{t-1}^* + \gamma^*(t-1)\right) + \Psi^* \Delta x_{t-1}^* + a_0^* + u_t^*, \qquad (7)$$

where no feedbacks, either short term or long term, is allowed from Canada to the US.

When the rank of the matrix Π^* is $r^* < m^*$, Π^* can be decomposed as $\Pi^* = \alpha^* \beta^{*\prime}$ where α^* is a $m^* \times r^*$ matrix of loading coefficients and $\beta^{*\prime}$ is a $r^* \times m^*$ matrix of cointegrating vectors.

3.3. The VECX*(r,r*) Model

After the models for Canada and the US are estimated, they are combined to form a complete econometric model, a VECX*(r,r*). In this resultant full model, all variables will be endogenous. The VECX*(r,r*) model can be written as

$$z_t = \Phi_1 z_{t-1} + \Phi_2 z_{t-2} + a + bt + H v_t, \qquad (8)$$

where

$$z_t = \begin{bmatrix} x_t \\ x_t^* \end{bmatrix}, a = \begin{bmatrix} a_0 - \Pi\gamma + \Lambda(a_0^* - \Pi^*\gamma^*) \\ a_0^* - \Pi^*\gamma^* \end{bmatrix}, b = \begin{bmatrix} \Pi\gamma + \Lambda\Pi^*\gamma^* \\ \Pi^*\gamma^* \end{bmatrix}, \Phi_1 = I + \widetilde{\Pi} + \widetilde{\Psi},$$

$$\Phi_2 = -\widetilde{\Psi}, \widetilde{\Pi} = \begin{bmatrix} \Pi + \Lambda[0 & \Pi^*] \\ 0 & \Pi^* \end{bmatrix}, \widetilde{\Psi} = \begin{bmatrix} \Psi + \Lambda[0 & \Psi^*] \\ 0 & \Psi^* \end{bmatrix}, H = \begin{bmatrix} I_m & \Lambda \\ 0 & I_{m^*} \end{bmatrix}, v_t = \begin{bmatrix} u_t \\ u_t^* \end{bmatrix}.$$

The covariance matrix $\Sigma = E(v_t v_t')$ can be freely estimated by the $\widetilde{m} \times \widetilde{m}$ matrix $\hat{\Sigma} = \sum_t \hat{v}_t \hat{v}_t' \,/\, T$ when no restriction is imposed on the covariance matrix.

This VECX*(r,r*) has 10 endogenous variables, $r + r^*$ cointegrating relations, and 10-$(r + r^*)$ stochastic trends. This model allows for a large degree of interdependence through two different channels. First, through the effects of x_t^* variables which are generally large; shocks to the US have considerable impacts on Canada. Second, through the covariances between errors, $\hat{\Sigma}$, which are expected to be small relative to the direct x_t^* effects.

3.4. Forecast Evaluation Criteria

In this exercise, "root mean squared forecast error" (RMSFE) is the criteria which will be used to evaluate forecasting results of different models. If z_{t+h} is the level of a variable and $\hat{z}(t+h,t)$ is the forecast of that variable formed at time t, then the forecast error for h-step ahead forecasts is given by:

$$e_t(h) = [z_{t+h} - \hat{z}(t+h,t)]. \qquad (9)$$

Having computed the forecast error, RMSFE is defined as

$$RMSFE(h,n) = 100\sqrt{n^{-1}\sum_{t=T}^{T+n-1} e_t^2(h)} \qquad (10)$$

where the forecast evaluation period is from T+1 to T+n.

4. Forecast Comparison

In this section, we estimate the models of Canada and the US assuming that cointegrating vectors are exactly identified. The cointegration ranks for Canada and the US are denoted by r and r*, respectively. In order to have a broad picture of the effects of cointegration rank restrictions on forecasting performance, we consider models with different cointegration ranks. We allow r to vary between zero and five, r=0,…,5, and r* to vary between zero and two, r*=0,1,2. Five and two are the maximum number of the long run relations which are suggested by economic theory, and also by trace and maximal eigenvalue tests, for the models of Canada and the US, respectively.[1] combining models of Canada and the US, we will have 18 different exactly identified VECX*(r,r*) specifications.

We conduct our forecasting exercise over short to medium forecast horizons (up to eight quarters ahead) since this is the relevant time horizon for central banks when setting monetary policy.

In order to derive reliable conclusions from this study, we had to base the analysis on a relatively large number of quarterly forecasts and therefore, we decided to generate (about) 50 quarterly forecasts for each forecast horizon. Hence, we estimate the cointegrated VARs of Canada and the US over the period of 1958Q1-1990Q4 and use data from 1991Q1 to 2004Q2 to evaluate the recursive out of sample forecasting performance. To do so, we estimate the models of Canada and the US over 1958Q1-1990Q4 (132 observations) separately, then combine them into a full system and compute one-to-eight quarters ahead forecasts, i.e. forecasts for 1991Q1-1992Q4. Then we extend the sample by one quarter at the end of the sample and re-estimate the models but this time over the period 1958Q1-1991Q1 and compute again a set of *1*-,*2*-,...,*8*-quarters ahead forecasts for the period 1991Q2-1993Q1. We continue this procedure until we reach 2004Q1. Evidently, at 2004Q1 we can produce only a set of *1*-quarter ahead forecasts. At the end we would have generated 54 sets of *1*-quarter ahead forecasts, 53 sets of *2*-quarters ahead forecasts, ..., and 47 sets of *8*-quarters ahead forecasts. Therefore, the RMSFEs rely on a different number of forecast errors for each horizon, ranging from 54 forecast errors for the *1*-quarter ahead forecasts to 47 for the *8*-quarters ahead forecasts.

---

1. Both the trace and maximal eigenvalue tests denote two cointegrating relations for the US. This is also consistent with what the theory suggests. Economic theory suggests that the term structure and Fisher equation hold in the US model. For the Canadian model, trace and maximal eigenvalue tests indicate the presence of three and five cointegrating relations, respectively. Based on the theory, the cointegration rank should be five since five long run relations, i.e. the term structure, Fisher equation, interest rate parity, purchasing power parity, and output convergence, are expected to hold in the Canadian model. For more details, see Barakchian (2012b).

We compare the RMSFEs of the exactly-identified VECX*(r,r*) models with those of VAR in first differences (DVAR), and VAR(1) and VAR(2) specifications in levels. It is worth mentioning that in our exercise, DVAR corresponds to the VECX*(0,0). Also, since we wanted to make the forecasts from the VAR in levels to be comparable to the forecasts from the VECX*(r,r*), we modeled the VAR in levels following the same procedure as the VECX*(r,r*): for VAR(1) model, first we estimate a VARX*(1) for Canada using the equation,

$$x_t = \Psi z_{t-1} + \Lambda x_t^* + a_0 + b_0 t + u_t,$$

where x_t and x_t^* are defined as before. Then a VAR(1) for the US is estimated by:

$$x_t^* = \Psi^* x_{t-1}^* + a_0^* + b_0^* t + u_t^*.$$

At the end, the models of the US and Canada are combined to form a full system required for forecasting.[1] The RMSFEs of Canadian inflation, output, short and long term interest rates, computed for the exactly-identified VECX*(r,r*) are reported in Tables 1 to 4. In order to have a better picture of the results, the best VECX* and VAR (DVAR) models, i.e. the models with the least RMSFE, are depicted in Table 5. It can be observed from these tables that:

1. To construct the VAR(2), Z_{t-2} is added to the model of Canada and X_{t-2}^* is added to the model of US.

Table 1: RMSFE of Canadian inflation in exactly-identified VECX*(r,r*) models of Canada and US

		Forecast Horizon							
		h=1	**h=2**	**h=3**	**h=4**	**h=5**	**h=6**	**h=7**	**h=8**
	#	54	53	52	51	50	49	48	47
r	**r***								
1	0	0.530	0.591	0.651	0.694	0.673	0.641	0.663	0.683
2	0	0.527	0.574	0.622	0.661	0.638	0.602	0.619	0.626
3	0	0.529	0.564	0.605	0.634	0.606	0.560	0.577	0.581
4	0	0.528	0.554	0.592	0.620	0.577	0.528	0.559	0.571
5	0	0.524	0.546	0.578	0.602	0.556	0.505	0.528	0.532
0	1	0.548	0.615	0.664	0.702	0.637	0.560	0.588	0.621
1	1	0.530	0.584	0.638	0.681	0.666	0.641	0.660	0.678
2	1	0.526	0.566	0.608	0.646	0.629	0.600	0.611	0.615
3	1	0.526	0.554	0.588	0.615	0.592	0.551	0.562	0.563
4	1	0.524	0.542	0.572	0.600	0.562	0.519	0.544	0.552
5	1	0.521	0.535	0.560	0.582	0.543	0.499	0.519	0.520
0	2	0.551	0.621	0.672	0.697	0.628	0.545	0.571	0.603
1	2	0.523	0.571	0.626	0.662	0.646	0.620	0.642	0.663
2	2	0.520	0.554	0.595	0.625	0.606	0.572	0.585	0.588
3	2	0.523	0.550	0.585	0.606	0.583	0.541	0.554	0.555
4	2	0.524	0.542	0.573	0.594	0.556	0.512	0.536	0.544
5	2	0.520	0.536	0.561	0.576	0.536	0.488	0.506	0.506
DVAR		0.553	0.624	0.677	0.715	0.645	0.563	0.595	0.629
VAR(1)		0.517	0.514	0.538	0.568	0.543	0.522	0.531	0.536
VAR(2)		0.521	0.536	0.561	0.579	0.543	0.501	0.518	0.516

Notes: Sequential out-of-sample forecasts from 1991Q1 to 2004Q2, estimation period 1958Q1 to 1990Q4. # indicates the number of point forecasts available to compute the RMSFE. r and r* are the cointegration ranks of the models of Canada and US in the VECX*(r, r*), respectively. DVAR is equivalent to VECX*(0, 0).

Table 2: RMSFE of Canadian output in exactly-identified VECX*(r,r*) models of Canada and US

		Forecast Horizon							
		h=1	**h=2**	**h=3**	**h=4**	**h=5**	**h=6**	**h=7**	**h=8**
	#	54	53	52	51	50	49	48	47
r	**r***								
1	0	0.617	1.000	1.378	1.732	2.050	2.269	2.441	2.583
2	0	0.594	0.922	1.262	1.597	1.920	2.155	2.336	2.483
3	0	0.653	1.027	1.422	1.822	2.218	2.517	2.765	2.965
4	0	0.682	1.086	1.540	2.009	2.453	2.803	3.068	3.262
5	0	0.674	1.067	1.524	2.009	2.481	2.878	3.201	3.469
0	1	0.581	0.914	1.276	1.625	1.938	2.152	2.314	2.441
1	1	0.610	0.987	1.361	1.717	2.043	2.271	2.451	2.598
2	1	0.587	0.910	1.245	1.582	1.915	2.161	2.352	2.507
3	1	0.648	1.017	1.404	1.801	2.197	2.501	2.752	2.958
4	1	0.680	1.088	1.548	2.030	2.492	2.864	3.150	3.365
5	1	0.676	1.078	1.548	2.051	2.547	2.973	3.326	3.625
0	2	0.609	1.034	1.502	1.958	2.366	2.662	2.884	3.045
1	2	0.644	1.131	1.623	2.094	2.519	2.834	3.076	3.260
2	2	0.615	1.039	1.474	1.902	2.303	2.604	2.828	2.995
3	2	0.672	1.123	1.593	2.064	2.513	2.850	3.113	3.312
4	2	0.706	1.193	1.725	2.266	2.762	3.148	3.428	3.627
5	2	0.695	1.160	1.676	2.208	2.708	3.115	3.431	3.684
DVAR		0.591	0.936	1.303	1.648	1.952	2.155	2.311	2.433
VAR(1)		0.620	1.035	1.483	1.969	2.443	2.854	3.189	3.468
VAR(2)		0.720	1.229	1.797	2.380	2.933	3.391	3.763	4.073

Notes: Sequential out-of-sample forecasts from 1991Q1 to 2004Q2, estimation period 1958Q1 to 1990Q4. # indicates the number of point forecasts available to compute the RMSFE. r and r* are the cointegration ranks of the models of Canada and US in the VECX*(r, r*), respectively. DVAR is equivalent to VECX*(0, 0).

Table 3: RMSFE of Canadian short term interest rate in exactly-identified VECX*(r,r*) models of Canada and US

		Forecast Horizon							
		h=1	**h=2**	**h=3**	**h=4**	**h=5**	**h=6**	**h=7**	**h=8**
	#	54	53	52	51	50	49	48	47
r	**r***								
1	0	0.182	0.287	0.350	0.431	0.491	0.540	0.583	0.598
2	0	0.180	0.282	0.338	0.414	0.476	0.528	0.574	0.591
3	0	0.186	0.294	0.356	0.437	0.507	0.566	0.622	0.647
4	0	0.191	0.304	0.366	0.444	0.508	0.561	0.612	0.639
5	0	0.193	0.310	0.375	0.454	0.518	0.574	0.625	0.652
0	1	0.177	0.269	0.322	0.404	0.466	0.514	0.559	0.568
1	1	0.183	0.285	0.345	0.424	0.483	0.530	0.573	0.587
2	1	0.181	0.280	0.333	0.406	0.467	0.516	0.562	0.579
3	1	0.186	0.293	0.352	0.432	0.500	0.557	0.612	0.638
4	1	0.191	0.303	0.366	0.445	0.510	0.563	0.614	0.641
5	1	0.192	0.308	0.375	0.454	0.520	0.576	0.628	0.656
0	2	0.177	0.268	0.321	0.404	0.467	0.517	0.564	0.573
1	2	0.183	0.283	0.340	0.418	0.478	0.526	0.571	0.585
2	2	0.181	0.279	0.329	0.403	0.465	0.518	0.567	0.585
3	2	0.186	0.289	0.343	0.420	0.487	0.545	0.600	0.624
4	2	0.190	0.297	0.348	0.418	0.477	0.528	0.579	0.605
5	2	0.192	0.301	0.354	0.424	0.484	0.539	0.591	0.617
DVAR		0.174	0.270	0.327	0.411	0.473	0.524	0.569	0.578
VAR(1)		0.188	0.290	0.358	0.430	0.481	0.522	0.55	0.552
VAR(2)		0.191	0.304	0.363	0.435	0.496	0.549	0.597	0.615

Notes: Sequential out-of-sample forecasts from 1991Q1 to 2004Q2, estimation period 1958Q1 to 1990Q4. # indicates the number of point forecasts available to compute the RMSFE. r and r* are the cointegration ranks of the models of Canada and US in the VECX*(r, r*), respectively. DVAR is equivalent to VECX*(0, 0).

Table 4: RMSFE of Canadian long term interest rate in exactly-identified VECX*(r,r*) models of Canada and US

		Forecast Horizon							
		h=1	**h=2**	**h=3**	**h=4**	**h=5**	**h=6**	**h=7**	**h=8**
	#	54	53	52	51	50	49	48	47
r	**r***								
1	0	0.104	0.163	0.203	0.247	0.276	0.300	0.334	0.356
2	0	0.103	0.162	0.201	0.245	0.274	0.298	0.332	0.354
3	0	0.104	0.165	0.206	0.250	0.280	0.304	0.338	0.360
4	0	0.107	0.174	0.220	0.266	0.295	0.319	0.352	0.375
5	0	0.108	0.177	0.225	0.271	0.303	0.328	0.363	0.387
0	1	0.098	0.146	0.183	0.231	0.263	0.287	0.328	0.352
1	1	0.104	0.166	0.208	0.253	0.284	0.312	0.351	0.377
2	1	0.104	0.164	0.206	0.251	0.282	0.309	0.347	0.373
3	1	0.105	0.169	0.213	0.258	0.290	0.318	0.356	0.382
4	1	0.108	0.179	0.228	0.275	0.307	0.334	0.372	0.399
5	1	0.109	0.181	0.231	0.280	0.313	0.343	0.381	0.409
0	2	0.100	0.151	0.190	0.240	0.273	0.302	0.344	0.372
1	2	0.107	0.173	0.218	0.265	0.299	0.330	0.371	0.401
2	2	0.108	0.173	0.219	0.267	0.302	0.335	0.377	0.407
3	2	0.110	0.179	0.227	0.275	0.309	0.341	0.381	0.411
4	2	0.114	0.192	0.247	0.298	0.333	0.364	0.403	0.432
5	2	0.115	0.195	0.252	0.305	0.342	0.375	0.415	0.445
DVAR		0.099	0.146	0.180	0.227	0.255	0.275	0.311	0.331
VAR(1)		0.109	0.170	0.219	0.265	0.292	0.313	0.335	0.347
VAR(2)		0.113	0.192	0.248	0.299	0.333	0.358	0.391	0.412

Notes: Sequential out-of-sample forecasts from 1991Q1 to 2004Q2, estimation period 1958Q1 to 1990Q4. # indicates the number of point forecasts available to compute the RMSFE. r and r* are the cointegration ranks of the models of Canada and US in the VECX*(r, r*), respectively. DVAR is equivalent to VECX*(0, 0).

Table 5: Models with the best forecasting performance

Forecast Horizon	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8
#	54	53	52	51	50	49	48	47
Inflation								
VECX*	2,2 5,2	5,1	5,1	5,2	5,2	5,2	5,2	5,2
RMSFE	0.520	0.535	0.560	0.576	0.536	0.488	0.506	0.506
	VAR(1)	VAR(1)	VAR(1)	VAR(1)	VAR(1) VAR(2)	VAR(2)	VAR(2)	VAR(2)
RMSFE	0.517	0.514	0.538	0.568	0.543	0.501	0.518	0.516
Ratio	1.006	1.041	1.041	1.014	0.987	0.974	0.977	0.981
Output								
VECX*	0,1	2,1	2,1	2,1	2,1	0,1	0,1	0,1
RMSFE	0.581	0.910	1.245	1.582	1.915	2.152	2.314	2.441
	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR
RMSFE	0.591	0.936	1.303	1.648	1.952	2.155	2.311	2.433
Ratio	0.983	0.972	0.955	0.960	0.981	0.999	1.001	1.003
Short term interest rate								
VECX*	0,1 0,2	0,2	0,2	2,2	2,2	0,1	0,1	0,1
RMSFE	0.177	0.268	0.321	0.403	0.465	0.514	0.559	0.568
	DVAR	DVAR	DVAR	DVAR	DVAR	VAR(1)	VAR(1)	VAR(1)
RMSFE	0.174	0.270	0.327	0.411	0.473	0.522	0.550	0.552
Ratio	1.017	0.993	0.982	0.981	0.983	0.985	1.016	1.029
Long term interest rate								
VECX*	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
RMSFE	0.098	0.146	0.183	0.231	0.263	0.287	0.328	0.352
	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR	DVAR
RMSFE	0.099	0.146	0.180	0.227	0.255	0.275	0.311	0.331
Ratio	0.990	1.000	1.017	1.018	1.031	1.044	1.055	1.063

Notes: Sequential out-of-sample forecasts from 1991Q1 to 2004Q2, estimation period 1958Q1 to 1990Q4. # indicates the number of point forecasts available to compute the RMSFE. For each variable, the model with the best forecasting performance among the VECX* models and also the best model out of DVAR, VAR(1) and VAR(2) models, with their RMSFEs are depicted. 'Ratio' corresponds to the ratio of the RMSFE of the best VECX* model to the RMSFE of the best VAR (or DVSR) model. DVAR is equivalent to VECX*(0, 0).

1. As expected, the RMSFEs of all variables monotonically increase, as forecast horizon increases (this pattern holds for all exactly-identified VECX*(r,r*) models). The RMSFEs of these variables at *8*-quarters ahead horizon are three to four times larger than the RMSFEs at 1-quarter ahead horizon. But this is not the case for inflation. The RMSFE of inflation peaks at 4-quarters ahead horizon and then it falls (this pattern also holds for all exactly-identified VECX*(r,r*) models). The RMSFE of inflation barely increases between 1- and 8-quarters ahead horizons and it even decreases for VECX*(5,2).
2. VAR(1) almost always beats VAR(2). It is only outperformed by VAR(2) when forecasting inflation at horizons 6-, 7-, and 8-quarters ahead.
3. DVAR has an inferior performance compared to VAR(1) and VAR(2) when forecasting inflation but has a superior one when forecasting the other variables (there is an exception: DVAR is outperformed by VAR(1) at the horizons of 6-, 7-, and 8-quarters ahead when forecasting short term interest rate).
4. There is not a single VECX* model which performs the best when forecasting either across variables or across horizons.
5. As the cointegration ranks (r, r*) imposed on the models are increased, the accuracy of the forecasts of output, short and long term interest rates typically decrease [compare VECX*(0,0) and VECX*(0,1) on the one hand with VECX*(4,2) and VECX*(5,2) on the other hand]. This result is similar to the findings of Lin and Tsay (1996) and contrasts the results of Clements and Hendry (1995).[1] But this pattern does not hold for the RMSFE of inflation.
6. As the forecast horizon gets longer, the VECX* model with the maximum number of cointegration ranks performs the best when forecasting inflation. On the opposite, the VECX* models with the minimum number of ranks do the best when forecasting the other variables [look at VECX*(0,1), and also note that at longer horizons DVAR, which corresponds to VECX*(0,0), performs even better than VECX*(0,1)].

1. Brandner and Kunst (1990) also find a similar result. In fact, because of the costs of imposing too much cointegration, they recommend the use of VAR in differences.

7. In Table 4, ‘Ratio’ corresponds to the ratio of the RMSFE of the best VECX* model to the RMSFE of the best VAR (or DVAR) model. As this measure shows, there is no significant difference between the performance of the best VECX* and best VAR (or DVAR) models, and the RMSFEs of these models are almost the same. Overall, we can conclude that a VAR model for forecasting inflation and a DVAR model for forecasting the rest of the variables work as well as the best VECX* model (and even better).

5. Conclusion

In this paper, we provided a comprehensive review over the subject of ‘cointegration and forecasting’. We also examined the effects of cointegration rank restrictions on forecasting performance of VAR models through conducting an empirical exercise in the framework of a new two-country Canada-US model. The results showed that as the cointegration ranks (r,r*) imposed on the models are increased, the accuracy of the forecasts (except for inflation) typically decrease. We also found that a VAR/DVAR model forecasts as well as the best cointegrated VAR model (and even better). Therefore, it seems that using cointegration techniques does not pay a dividend.

References

- Anderson, R.G., D.L. Hoffman and R.H. Rasche (2002). " A Vector Error-correction Forecasting Model of the US Economy", *Journal of Macroeconomics*, 24, 569-598.
- Assenmacher-Wesche, K., and M.H. Pesaran (2008). " Forecasting the Swiss Economy Using VECX* Models: An Exercise in Forecast Combination across Models and Observation Windows", *National Institute Economic Review*, 203, 91-108.
- Barakchian, S.M. (2012a). " Do Long Run Theory Restrictions Help in Forecasting?", *Journal of Forecasting*, 31, 401-422.
- Barakchian, S.M. (2012b). " Transmission of US Monetary Policy into the Canadian Economy: A Structural Cointegration Analysis", Cambridge University, Mimeo.
- Bernanke, B.S., and I. Mihov (1998). " Measuring Monetary Policy", *Quarterly Journal of Economics*, 113, 869-902.
- Brandner, P., and R.M. Kunst (1990). " Forecasting Vector Autoregressions-The Influence of Cointegration: A Monte Carlo Study", *Research Memorandum 265*, Institute for Advanced Studies, Vienna.
- Christiano, L.J., M. Eichenbaum and C.L. Evans (1996). "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds", *Review of Economics and Statistics*, 78, 16-34.
- Christiano, L.J., M. Eichenbaum and C.L. Evans (1999). "Monetary Policy Shocks: What Have We Learned and to What End?", in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1, Amsterdam, Elsevier, 65-148.
- Christoffersen, P.F., and F.X. Diebold (1998). "Cointegration and Long-Horizon Forecasting", *Journal of Business and Economic Statistics*,16, 450-458.
- Clements, M.P., and D.F. Hendry (1993). "On the Limitations of Comparing Mean Squared Forecast Errors", *Journal of Forecasting*, 12, 617-637.

– Clements, M.P., and D.F. Hendry (1995). "Forecasting in Cointegrated Systems", *Journal of Applied Econometrics*, 10, 127-146.

– Clements, M.P., and D.F. Hendry (1996). " Intercept Corrections and Structural Change", *Journal of Applied Econometrics*, 11, 475-494

– Clements, M.P., and D.F. Hendry (1998). *Forecasting Economic Time Series*, Cambridge, Cambridge University Press.

– Clements, M.P., and D.F. Hendry (1999). *Forecasting Non-stationary Economic Time Series*, Cambridge, MIT Press.

– Diebold, F.X. (1998). "The Past, Present, and Future of Macroeconomic Forecasting", *Journal of Economic Perspectives,* 12, 175-92.

– Diebold, F.X., G.D. Rudebusch and S.B. Aruoba (2006). "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach", *Journal of Econometrics*, 131, 309-338.

– Engle, R.F., and C.W.J. Granger (1987). "Cointegration and Error Correction Representation: Estimation and Testing", *Econometrica,* 55, 251-276.

– Engle, R.F., and B.S. Yoo (1987). "Forecasting and Testing in Cointegrated Systems", *Journal of Econometrics*, 35, 143-159

– Estrella, A., and F.S. Mishkin (1998). "Predicting U.S. Recessions: Financial Variables as Leading Indicators", *Review of Economics and Statistics*, 80, 45-61.

– Hendry, D.F. (2006). "Robustifying Forecasts from Equilibrium-correction Models", *Journal of Econometrics*, 135, 399-426.

– Hendry, D.F., and M.P. Clements (2003). "Economic Forecasting: Some Lessons from Recent Research", *Economic Modelling*, 20, 301-329.

– Hoffman, D.L., and R.H. Rasche (1996). "Assessing Forecast Performance in a Cointegrated System", *Journal of Applied Econometrics*, 11, 495-517.

– Jacobson, T., P. Jansson, A. Vredin and A. Warne (2001). "Monetary Policy Analysis and Inflation Targeting in a Small Open Economy: A VAR

Approach”, *Journal of Applied Econometrics,* 16, 487-520.

- Johansen, S. (1988). “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dynamics and Control,* 12, 231-254.
- Johansen, S. (1991). “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”, *Econometrica,* 59, 1551-1580.
- Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford, Oxford University Press.
- Lastrapes, W.D. (2002). “Comments on“ "A Vector Error-correction Forecasting Model of the US Economy”, *Journal of Macroeconomics,* 24, 607-611.
- Lin, J.-L., and R.S. Tsay (1996). “Cointegration Constraint and Forecasting: An Empirical Examination”, *Journal of Applied Econometrics*, 11, 519-538.
- Lucas, R.E. (1976). “Econometric Policy Evaluation: A Critique”, *Carnegie Rochester Conference Series on Public Policy,* 1, 19–46.
- McCrae, M., Y. Lin, D. Pavlik, and C.M. Gulati (2002). “ Can Cointegration-based Forecasting Outperform Univariate Models? An Application to Asian Exchange Rates”, *Journal of Forecasting,* 21, 355-380.
- Pesaran, M.H., Y. Shin and R.J. Smith (2000). “Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables”, *Journal of Econometrics*, 97, 293-343.
- Pesaran, M.H., T. Schuermann, and V. Smith (2009). “Forecasting Economic and Financial Variables with Global VARs”, with discussions and rejoinder, *International Journal of Forecasting,* 25, 642-675
- Sims, C.A. (1980). “Macroeconomics and Reality”, *Econometrica,* 48, 1-48.
- Sims, C.A., and T. Zha (2006). “Does Monetary Policy Generate Recessions?”, *Macroeconomic Dynamics*, 10, 231-272.

– Söderlind, P., and A. Vredin (1996). “Applied Cointegration Analysis in the Mirror of Macroeconomic Theory”, *Journal of Applied Econometrics,* 11, 363-381.

– Swanson, N.R. (2002). “Comments on ‘A Vector Error-correction Forecasting Model of the US Economy’”, *Journal of Macroeconomics,* 24, 599-606.

– Wang, Z., and D.A. Bessler (2004). “Forecasting Performance of Multivariate Time Series Models with a Full and Reduced Rank: An Empirical Examination”, *International Journal of Forecasting*, 20, 683-695.

Appendix

Definitions and Sources of Variables

The data set contains quarterly observations on the US and Canada, from 1958Q1 to 2004Q2. The Canadian variables included are (log) real per capita output, y_t, (log) price level, p_t, the nominal quarterly short term interest rate, r_t^s, the nominal quarterly long term interest rate, r_t^l, and (log) exchange rate, e_t. Specifically

$$y_t = ln\left(\frac{GDP_t}{POP_t}\right), p_t = ln(P_t)\,, e_t = ln(E_t)\,, r_t^s = 0.25\, ln\left(1 + \left(\frac{R_t^s}{100}\right)\right),$$

$$\mathrm{r_t^l} = 0.25 \ln\left(1 + \left(\frac{\mathrm{R_t^l}}{100}\right)\right),$$

where GDP_t is real gross domestic product volume index (seasonally adjusted and indexed at 100 in 2000), P_t is the consumer price index (seasonally adjusted and indexed at 100 in 2000), R_t^s is the treasury bill rate (percent per annum), R_t^l is the 10 years bond rate (percent per annum), and E_t is the Canadian currency per US dollar (indexed at 100 in 2000). The US variables, $y_t^*, p_t^*, {r_t^s}^*, {r_t^l}^*$ are constructed using the same method.

POP_t is constructed as a quarterly series through linear interpolation of the annual series and then converted into an index number (indexed at 1 in 2000). The oil price variable, p_t^o, is constructed as $p_t^o = \ln(POIL_t)$, where $POIL_t$ is the average price of crude oil in terms of US dollar per barrel (indexed at 100 in 2000).

The data were obtained from the International Financial Statistics, IMF. GDP and CPI series are seasonally adjusted using the X12-ARIMA method.