Some Advances in Restricted Forecasting Theory for Multiple Time Series

by

Nicolás Gómez Castillo

A dissertation submitted to the IDEA program in partial fulfillment of the requirements for the Degree

Doctor in Economy

Universitat Autònoma de Barcelona
Departament d'Economia i d'Historia Econòmica
International Doctorate in Economic Analysis
Bellaterra, Barcelona

2007

"*Some Advances in Restricted Forecasting Theory for Multiple Time Series*", a dissertation prepared by Nicolás Gómez Castillo in partial fulfillment of the requirements for the degree, Doctor in Economy.

Dr. Víctor M. Guerrero
Instituto Tecnológico Autónomo de México
Thesis Advisor

Dr. Michael Creel
Universitat Autònoma de Barcelona
Thesis Tutor

DEDICATION

"Este trabajo se lo dedico a la memoria de mi madre, Conchita. Siempre has estado presente en mi recuerdo con tu ejemplo para darme fuerza y coraje en los momentos difíciles".

ACKNOWLEDGMENTS

First I would like to express my gratitude to my mentor, Dr. Víctor M. Guerrero, who introduced me to the subject of Forecasting Economic Time Series, passing onto me all his passion and interests. This thesis would not have been written without his incessant support and motivation. I want to thank Dr. Michael Creel for his useful comments and assistance on this thesis. Three anonymous referees and the ex-editor of the International Journal of Forecasting, Prof. Jan De Gooijer, provided insightful comments and suggestions that helped to enhance the quality of my work which derived in the publication of Chapter 2. I am grateful to Eva Senra for her valuable comments and suggestions that resulted in a substantial improvement of this work. I am also indebted to an anonymous evaluator of the all thesis for his/her helpful comments and opinions of this work, thank you.

I am in debt with Departament d'Economia i d'Historia Econòmica for give me the opportunity to do my doctoral studies, to know some marvelous people and to be in this beautiful city. Thanks to the Agencia Española de Cooperación Internacional for financial support during my stance in Barcelona.

I would like to thank all who contributed to this thesis by their discussions, comments and suggestions or by simply being where I needed them to be. In particular Dr. Jorge de la Vega and Dr. Raymundo Peralta for pushing me when I was static in the saddle point.

Finally, I thank Marbella, my wife, for her constant encouragement and patient through out the years of my research, without what this thesis would not be possible.

CONTENTS

LIST OF TABLES vii

1 GENERAL INTRODUCTION 1

1.1 Forecasting 1

1.2 Restricted Forecasting 5

1.3 Contribution 7

1.3.1 Chapter 2 7

1.3.2 Chapter 3 9

1.3.3 Chapter 4 10

1.4 Further extensions 11

2 RESTRICTED FORECASTING WITH VAR MODELS: AN ANALYSIS OF A TEST FOR JOINT COMPATIBILITY BETWEEN RESTRICTIONS AND FORECASTS 16

2.1 Introduction 16

2.2 Methodology 19

2.2.1 The VEC Model 20

2.2.2 Stochastic linear restrictions on future values 23

2.2.3 Restricted forecasts 25

2.2.4 Compatibility test 26

2.3 Additivity of the compatibility test 27

2.3.1 Decomposition of the JCT into single tests 27

2.3.2 Decomposition of **K** 29

2.4 A feasible JCT for estimated processes 32

2.4.1 A Monte Carlo study of the JCT 33

2.4.2 The VAR forecast for estimated processes 37

2.4.2.1 Restricted forecast with an estimated process . . 39

2.4.2.2 The feasible JCT 39

2.4.3 A Monte Carlo study for the feasible JCT 41

2.4.4 A simulation study for the Mexican data example 43

2.5 Empirical illustration 44

2.5.1 Order of integration 46

2.5.2 VEC estimation 47

2.5.3 Forecasts restricted to fulfill targets 51

2.5.4 Forecasts restricted by unrealistic targets 57

2.6 Conclusions 58

2.7 Appendices 60

2.7.1 Appendix A. Approximation of MSE($\mathbf{Y}_{F,H}$) for an estimated process 60

2.7.2 Appendix B. Restricted forecasts for estimated process . . 62

3 RESTRICTED VAR FORECASTS THAT TAKE INTO ACCOUNT AN

EXPECTED STRUCTURAL CHANGE 68

3.1 Introduction 68

3.2 Methodology 71

3.3 VAR with structural change 73

3.3.1 Deterministic changes 74

3.3.2 Stochastic changes 78

3.3.3 The forecast 79

3.4 Restricted forecasts of a VAR with structural change 81

3.5 Empirical illustration 83

3.5.1 The data 83

3.5.2 Order of integration 85

3.5.3 VEC estimation 86

3.5.4 Restricted forecasts of the Mexican economy 90

3.5.4.1 Scenario 1, deterministic innovational change . . 93

3.5.4.2 Scenario 2, variance innovational change 99

3.6 Conclusions 103

3.7 Appendix. Shape of $\mathbf{K}(a)$ 104

4 RESTRICTED VAR FORECASTS OF ECONOMIC TIME SERIES WITH CONTEMPORANEOUS CONSTRAINTS 108

4.1 Introduction 108

4.2 Preliminaries 111

4.2.1 Models 112
4.2.2 Forecasts 113
4.3 Restricted forecasts for contemporaneously constrained VAR models 116
4.4 An empirical illustration with the balance of payments account . . 121
4.4.1 Order of integration 122
4.4.2 Model estimation 122
4.4.2.1 Deficit variable 123
4.4.2.2 Income-expenditure system 123
4.4.3 Restricted forecasts 127
4.4.3.1 Univariate restricted forecast of deficit 127
4.4.3.2 Income-expenditure system, one-stage restricted forecast 127
4.4.3.3 Income-expenditure system, two-stage restricted forecast 128
4.5 An extension to cointegrated systems 130
4.6 Conclusions 140

LIST OF TABLES

2.1 p_α values for the compatibility test with binding restrictions . . . 36

2.2 p_α values for the compatibility test with unbinding restrictions . . 37

2.3 p_α values for the feasible compatibility test with binding restrictions 41

2.4 p_α values for the feasible compatibility test with unbinding restrictions 42

2.5 p_α values for the estimated Mexican system 43

2.6 ADF unit root test results 46

2.7 Johansen cointegration analysis 50

2.8 Feasible compatibility tests for Mexican economic targets 54

2.9 Forecasts of the Mexican system 57

2.10 Feasible compatibility tests for unrealistic economic targets . . . 59

3.1 ADF unit root tets results 86

3.2 Johansen cointegration analysis 89

3.3 VEC forecasts results for the Mexican system 98

3.4 Restricted forecasts with stochastic change for the Mexican system, $\widehat{\mathbf{X}}_{F,V}$ 103

4.1 Unit root ADF results 123

4.2 Johansen cointegration analysis 125

4.3 Forecast precision measures: model-based and standard MSEs . . 136

4.4 Forecast precision measures that take contegration into account . 137

CHAPTER 1

GENERAL INTRODUCTION

1.1 Forecasting

Forecasts of economic variables play an important role in business decision-making, government policy analysis, and economic research. For example, questions related to the capital structure policy upon the growth rate of a firm, the effects of changes in the monetary policies for the control of inflation upon productivity and employment, the impact of legislation about retirement income upon saving behavior. This work is an attempt to contribute, using classical and popular time series methods, to the multivariate forecasting methodology when additional future information is available.

Forecasts often are model-based, in the sense that the forecasting operation is carried out after the statistical identification and estimation of a suitable stochastic model that employs the available time series data. The forecasting procedures themselves are simply devices for utilizing the estimated model to project its output into the future in stochastic terms, normally through the evolution of the mean and the standard error bands associated with the forecast distribution. Some of the earliest and simplest univariate forecasting methods are the

Exponential Smoothing methods introduced by Holt (1957) and Winters (1960). Different versions of these forecasting procedures are based on the construction of forecast functions that depend on discounted past observations. However, these methods are *ad hoc* in the sense that they are implemented without respect to a suitable statistical model. These *ad hoc* methods can be justified nowadays by the univariate Auto-Regressive Integrated Moving Average (ARIMA) models. The popularization of the ARIMA model and its associated forecasting methodology was a result of the publication of Box and Jenkins' most influential book Time-Series Analysis, Forecasting and Control (1970, 1976). The success of these models lies in that they give the analyst a class of linear time series models which is sufficiently easy to be employed in practice and is general enough to provide a good representation of the data. In addition, ARIMA models have proved to be very successful at forecasting as compared to other univariate linear models. The ARIMA models will be used as a part of a two-stage forecast method in Chapter 4.

Quenouille (1957) and Tiao and Box (1981) introduced some important extensions of the univariate models to a multivariate framework in attempts to describe the dynamic relationship between the variables into consideration: the Vector Auto-Regressive (VAR) and the Vector Auto-Regressive Moving Average (VARMA) models. The use of these models require stationary stable time series systems. However, trends and variance fluctuations are quite common in practice.

Differencing and nonlinear transformations are methods introduced by Box and Jenkins (1970, 1976) for removing nonstationarities in the mean and variance of time series respectively. Since any VARMA model can be expressed as an approximately equivalent finite VAR model, and the estimation methods for the VAR model are simpler (equation-by-equation using linear Least Squares methods), it is preferred to use the latter. Due to the simple structure and linearity of the VAR model, it enjoys enormous popularity in empirical macroeconomic research and forecasting since Sims (1980) suggested its use as an alternative to classical multi-equation macroeconomic models. See Lütkepool (1991) for a good exposition of VAR and VARMA models and Clements and Hendry (2004) Chapter 4 for a description of several statistical forecasting models. An empirical illustration in Chapter 4 for a bivariate system with Mexican economic variables uses the VAR model.

Many economic time series have a trend behavior over time that makes them non-stationary, but groups of these time series variables may tend to drift together. The idea that non-stationary time series may keep together in the long-run is captured by the concept of cointegration which has several important implications in multiple time series analysis, as indicated by Engle and Granger (1987). The Vector Error Correction (VEC) models, which introduce the long-run relationship as a restriction into the VAR models, have proved to be useful when modeling and forecasting these kind of variables. Johansen (1995) provides likelihood-ratio tests for specifying the cointegration rank, which has become a standard technique nowadays. The methodological illustrations in Chapters 2 and 3 use a cointegrated system for the Mexican economy, where VEC models are required. A Monte Carlo simulation in Chapter 4 makes also use of this representation.

The model-based forecasting ability is crucial not only to determine the precision of the forecast but also to judge the adequacy of the model. There exist a number of studies in which a time series model seems to fit well in-sample, but perform poorly at obtaining out-of-sample predictions. Forecast comparisons are often based on the trace of the Mean Square Error (MSE) matrix. However, there are some works (e.g. Clements and Hendry, 1993) that criticized the use of the standard MSE measure because it is not invariant to non-singular, scale-preserving linear transformations. Lin and Tsay (1996) used the square root of the trace of the covariance matrix of out-of-sample forecasts errors as the main criterion and found that cointegration does not improve the forecast precision. Christoffersen and Diebold (1998) dealt with forecasting cointegrated variables and showed that nothing is lost by ignoring cointegration when forecasts are evaluated using the trace of the MSE, since such a measure fails to value the long-run forecasts. Thus, they suggested to use two MSE measures of forecasting performance. The first one is the trace MSE of the cointegrating combinations of the forecast errors. While the second one corresponds to a triangular representation of the cointegrated system that incorporates both the standard MSE and that of the aforementioned cointegrating combinations. These measures will be used to quantify the forecast precision in this work.

1.2 Restricted Forecasting

When forecasting time series variables, it is usual to use only the information provided by past observations to foresee potential future developments. However, if available, additional information should be taken into account to get the forecast. For example, let us consider a case where the Government announces an economic target for next year. Since the Government has the empowerment to implement the economic or social policies to approach the target, an analyst that does not consider this information to get the forecast and makes use only of the historical record of the variables, will not anticipate the change on the economic system. In fact, in a very influential article, Lucas (1976) established that predictions based on historical data would be invalid when a policy change affects the economy, since the economic agents are forward rather than backward-looking and adapt their expectations and behavior to the new policy stance. Thus, given some targets for the variables under study it is important to know the simultaneous future path that will lead to achieving those targets.

This work considers the case in which a system of variables are to be forecasted with the aid of a VAR model with a cointegration relationship. The paths projected forward into the future as a combination of the model-based forecasts and the additional information provides what is known as a restricted forecast.

When working with a single economic variable whose data consist of a univariate time series, the problem of incorporating the additional information has already been treated in the literature. Guerrero (1989) and Trabelsi and Hillmer (1989) obtained the corresponding optimum forecast, in minimum Mean Square Error (MSE) sense. Such a forecast allows an analyst to incorporate additional information as binding or unbinding restrictions.

The literature on restricted forecasts for multiple time series includes several papers. Doan, Litterman and Sims (1984), Green, Howrey and Hymans (1986), van der Knopp (1987) and Pankratz (1989) dealt with the combination of historical information with additional information provided by way of linear restrictions. Guerrero and Peña (2003) provided general results for the problem of combining data from two different sources of information in order to improve the efficiency of predictors. Some time series problems such as forecast updating when new information is available, forecast combination, interpolation and missing value estimation, among others, can be treated with their proposal. Pandher (2002) attacked the problem of modeling and forecasting a contemporaneously constrained system of time series within the state-space framework. Guerrero, Pena, Senra and Alegría (2005) focussed on implementing a Vector Error Correction (VEC) model for monitoring Mexican economic targets and verifying the compatibility between historical and additional information. A Joint Compatibility Test (JCT) for the restriction was proposed to that end.

1.3 Contribution

This dissertation is an attempt to contribute to the literature on *Restricted Forecasting Theory for Multiple Time Series within the VAR framework*. Specifically, Chapter 2 decomposes the JCT into single tests by a variance-covariance matrix associated with the restrictions and derives the formulas of a feasible JCT that accounts for estimated parameters. Chapter 3 develops, by Lagrangian optimization, the restricted forecasts of the multiple time series process with structural change, as well as its mean squared error. In addition, the univariate time series types of change presented by Tsay (1988) are considered here in a multivariate setting. Finally, Chapter 4 derives a methodology for forecasting multivariate time series that satisfy a contemporaneous binding constraint for which there exists a future target. A Monte Carlo study of a VEC model with one unit root shows that, for a forecast horizon large enough, the forecasts obtained with the proposed methodology are more efficient. A more detailed account of these contributions is provided below.

1.3.1 Chapter 2

This chapter presents a paper entitled *Restricted forecasting with VAR models: an analysis of a test for joint compatibility between restrictions and forecasts.* This work proposes a decomposition of the JCT for VEC models. It is shown that a variance-covariance matrix associated with the restrictions can be used to cancel out model dynamics and interactions between restrictions. This allows an analyst to interpret the joint compatibility test as a composition of the corresponding single restriction compatibility tests. These tests are useful to appreciate the contribution of each and every restriction to the joint compatibility between the whole set of restrictions and the unrestricted forecasts.

Since the JCT introduced in Guerrero *et al.* (2005) is based on asymptotic theory, a feasible version of this test was here derived. This new test also takes into account the parameter estimation of the model. A comparison between the JCT and the feasible JCT with the nominal test values, shows that the latter turns out to have better performance than the original one.

The proposed methodology was illustrated with a six-variable Mexican economic system with quarterly data. Implementation of the model focused on the economic targets for GDP, inflation rate and trade balance deficit for 2003, see SHCP (2002). A numerical simulation of this system was carried out to validate the use of the feasible JCT in this situation. It turned out that the economic targets were compatible with the unrestricted forecasts. Then, some unrealistic targets were considered to illustrate an incompatibility situation.

1.3.2 Chapter 3

Structural changes are commonly encountered in time series data analysis and the presence of those extraordinary events in the forecast horizon could easily mislead a time series model and its forecasts, thus resulting in erroneous conclusions. Thus, Chapter 3, called *Restricted VAR forecasts that take into account an expected structural change,* attacks the problem of forecasting when a structural change is expected to occur on an economic system during the forecast horizon. Both the deterministic and the stochastic structure of a VEC model are assumed to be affected by the structural change. The available information about the structural change is provided by a set of linear restrictions imposed on the future values of the variables involved. The restricted forecasts of the multiple time series process with structural change, as well as its mean squared error are here derived by Lagrangian optimization. These results generalize those of the univariate case obtained by Guerrero (1991). Furthermore, the univariate time series types of change presented by Tsay (1988) are considered here in a multivariate setting.

Nowadays, in Mexico, economic as well as political efforts concentrate on achieving the necessary consensus to advance in the fiscal, electric and pensions reforms. That is why this chapter presents an empirical illustration that makes use of Mexican macroeconomic data on six variables. The restricted forecasts take into account the economic targets for year 2004 announced by the Mexican Government (see, SHCP 2003) and an economic reform assumed for year 2005. It is also assumed that the economic reform will modify either the deterministic or the stochastic part of a VAR model and that its effect will initially impact GDP and prices.

1.3.3 Chapter 4

The last chapter is devoted to the paper *Restricted VAR forecasts of economic time series with contemporaneous constraints.* Here, a methodology is derived within the VAR framework, for forecasting multivariate time series that satisfy a contemporaneous binding constraint for which there exists a future target. Two ways of computing the restricted forecasts are proposed. The first one introduces the target as a linear restriction on the future values of the system, while the other introduces the target in the forecast of the aggregated variable (the contemporaneous binding constraint of the time series vector) which, in turn, is introduced as a restriction for the system forecasts. This methodology has natural applications when forecasting macroeconomic and financial time series that must satisfy accounting constraints, which are binding. The methodology is illustrated with the contemporaneously constrained income-expenditure system of the balance of payments account for the Mexican economy. Here the deficit (income minus expenditure) is the contemporaneous binding constraint for which a future restriction is given (i.e. the Government economic target, see SHCP 2004).

Since a cointegration relationship can be viewed as an unbinding contemporaneous constraint, the proposed methodology has immediate implications when forecasting cointegrated systems. The case of one unit root is considered in detail. As a starting point, it is required that the equilibrium of the system be reached at some finite point in the future. Such a restriction was imposed by letting the error correction term be equal to zero at that point. A Monte Carlo simulation of a VEC model with one unit root was carried out to compare the behavior of the unrestricted forecasts against those of the cointegrating restricted forecasts. Since the standard MSE has been severely criticized as a precision measure by several researches, three different precision measures are used in the comparisons.

1.4 Further extensions

Nowadays, Ministry Treasury in Mexico (the government agency engaged in tax collection and income distribution) faces a low tax collection problem. Historically, in Mexico the tax collection as a percentage of GDP is about 17% while other developed countries as US receives the 30%, Canada 39% and Germany 27%. Of course, the low tax collection limits the impact of government policies to deal with poverty, technological development, social security, energy, etc... For this reason a fiscal reform, among others, as been considered necessary to consolidate macroeconomic stability. However, some political forces have put up a lot of resistance to government proposals due to these are based mainly on the value-added tax. The restricted forecast methodology derived in Chapter 3 that account for structural changes could be used to construct scenarios of the Mexican economy with the main government proposals. This could contribute to judge the value-add tax controversy.

In order to control the levels of stocks and production of bank notes the Bank of Mexico (the Central Bank) needs to forecast the bank note circulation and measure their average life. To do that Bank of Mexico is supplied with a multi-anual monetary account (M1) time path forecast (the restriction). Nowadays, the M1 is disaggregated with historical weights and the forecast of each bank note is obtained by intervention analysis in the univariate ARMA setup. Therefore, modeling the bank notes system with an intervention analysis in a contemporaneously constrained VAR framework and using the restricted forecast methodology described in Chapter 4 could derive in a more precise bank note multivariate forecasts.

Consider the case where we are interested in getting the forecast of a time series with its unobserved components. For example, we decompose GDP into trend and cycle. Since the GDP variable is equal to the sum of its components, we would like to count on forecasts with the property that the forecast of GDP will be equal to the sum of the trend and cycle forecasts. However, with standard forecast methodologies, this adding-up property does not hold. In principle, a forecast methodology like that derived in Chapter 4 could help us with the problem of match the forecasts of a decomposed time series.

BIBLIOGRAPHY

[1] Box, G. E. P. & Jenkins, G. C. (1970, 1976). *Time Series Analysis: Forecasting and Control*, 1st and 2nd editions, San Francisco: Holden-Day.

[2] Clements, M. P. & Hendry, D. F. (2004). *A Companion to Economic Forecasting*, Oxford: Blackwell Publishing.

[3] Christoffersen, P. F. & Diebold, F. X. (1998). Cointegration and long-horizon forecasting. *Journal of Business & Economic Statistics* 16, 450-458.

[4] Doan, T., Litterman, R. & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions, *Econometric Reviews,* 3, 1-100.

[5] Engle, R. F. & Granger, C. W. J. (1987). Co-integration and error correction representation, estimation and testing, *Econometrica,* 55, 251-276.

[6] Greene, M. N., Howrey, E. P. & Hymans, S. W. (1986). The use of outside information in econometric forecasting, *in Model Reliability, eds. E. Huh and D. A. Belsley,* Cambridge, MA: MIT Press, 90-116.

[7] Guerrero, V. M. (1989). Optimal conditional ARIMA forecasts, *Journal of Forecasting,* 8, 215-229.

[8] Guerrero, V. M. & Peña D. (2003). Combining multiple time series predictors: a useful inferential procedure, *Journal of Statistical Planning and Inference,* 116, 249-276.

[9] Guerrero, V. M., Pena, B., Senra, E. & Alegría, A. (2005). Restricted forecasting with a VEC model: Validating the feasibility of economic targets, *Available in Eva Senra's website http://www2.uah.es/ .*

[10] Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages, *ONR Research Memorandum 52. Pittsburgh: Carnegie Institute of Technology*

[11] Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models,* Oxford University Press.

[12] Lin, J. L. & Tsay, R. S. (1996). Co-integration constraint and forecasting: an empirical examination. *Journal of Applied Econometrics* 11, 519-538.

[13] Lucas, R. E. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19-46.

[14] Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.

[15] Pandher, G. S. (2002). Forecasting multivariate time series with linear restrictions using constrained structural state-space models, *Journal of Forecasting,* 21, 281-300.

[16] Pankratz, A. (1989). Time series forecasts and extra-model information, *Journal of Forecasting,* 8, 75-83.

[17] Quenouille, M. H. (1957). *The Analysis of Multiple Time Series*, London: Griffin.

[18] Sims, C. A. (1980). Macroeconomics and reality, *Econometrica,* 48, 1-48.

[19] Tiao, G. C. & Box, G. E. P. (1981). Modeling multiple time series with applications, *Journal of the American Statistical Association,* 76, 802-816.

[20] Trabelsi, A. & Hillmer, S. C. (1989). A benchmarking approach to forecast combination, *Journal of Business and Economic Statistics,* 7, 353 - 362.

[21] Tsay, R. S. (1988). Outliers, level shifts, and variance changes in time series. *Journal of Forecasting* 7, 1 - 20.

[22] SHCP (2002). *Criterios Generales de Política Económica 2003*, México: Secretaría de Hacienda y Crédito Público.

[23] SHCP (2003). *Criterios Generales de Política Económica para la iniciativa de ley de ingresos y el proyecto de presupuesto de egresos de la federación correspondiente al año 2004*. México: Secretaría de Hacienda y Crédito Público.

[24] SHCP (2004). *Criterios Generales de Política Económica 2005.* México: Secretaría de Hacienda y Crédito Público.

[25] van der Knoop, H. S. (1987). Conditional forecasting with a multivariate time series model. *Economics Letters,* 22, 233 - 236.

[26] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages, *Management Science,* 6, 324 - 342.

CHAPTER 2

RESTRICTED FORECASTING WITH VAR MODELS: AN ANALYSIS OF A TEST FOR JOINT COMPATIBILITY BETWEEN RESTRICTIONS AND FORECASTS

2.1 Introduction

Decision making is mainly based on predictions of the most relevant variables for the problem at hand and time series data are usually employed to get forecasts. In the present case we let y_t denote the value of the variable at time t, then a forecast of the variable at time $T+h$ with information up to T is given by

$$\widehat{y}_{T+h} = f\left(y_T, y_{T-1}, \ldots\right).$$

If available, additional information should also be considered to get the forecast. For example, let us consider a case where the Government announces an economic target for next year. Since the Government has the empowerment to implement the economic or social policies to approach the target, the future information available (the target in this case) should be taken into account to get the forecast. Thus, given some targets for the variable under study it is important to know the future path that will lead to achieving that target. The path obtained as a combination of the forecasts from a model and the additional information provided by the targets (known as restricted forecasts) produce such a scenario of future values. Let us suppose that θ_{T+i} is the target for time $T+i$, thus the forecast will be given by

$$\widetilde{y}_{T+h} = f\left(\theta_{T+i}, y_T, y_{T-1}, \ldots\right).$$

When working with a single economic variable whose data consist of a univariate time series, the problem of incorporating the additional information has already been treated in the literature. Guerrero (1989) and Trabelsi and Hillmer (1989) obtained the corresponding optimum forecast, in minimum Mean Square Error (MSE) sense. Such a forecast allows an analyst to incorporate additional information as binding or unbinding restrictions.

The literature on restricted forecasts for multiple time series includes several papers. Doan, Litterman and Sims (1984), Green, Howrey and Hymans (1986), van der Knopp (1987) and Pankratz (1989) dealt with the combination of historical information with additional information provided by way of linear restrictions. Pandher (2002) attacked the same problem within the state-space framework. Guerrero, Pena, Senra and Alegría (2005) focussed on implementing a Vector Error Correction (VEC) model for monitoring Mexican economic targets and verifying the compatibility between historical and additional information. A joint compatibility test (JCT) for the restriction was proposed to that end.

When the targets impose linearly independent restrictions on the forecasts we can always obtain the corresponding restricted forecasts, no matter how discordant the targets are from the historical information. So, the JCT proves to be a useful tool to decide whether or not the targets are compatible with the forecasts based only on the historical record. Knowing that these two sources of information are compatible or not is useful in many ways. On the one hand, when the two sources of information are compatible with each other, incorporating the targets as additional information into the forecast will reduce the forecast MSE. On the other, when the test rejects compatibility, we are led to analyze the reason why this occurred. Then we would like to know which restrictions were the most likely causes of rejection. However, the JCT does not provide any clue about this. Therefore the need of being able to appreciate the individual contribution of each restriction to the joint compatibility arises. Here we address this issue by means of a decomposition of the JCT into tests for each individual restriction involved. Since the JCT introduced in Guerrero *et al.* (2005) is based on asymptotic theory, a feasible version of this test that works well with estimated parameters is here derived.

This article is organized as follows. Section 2.2 presents the statistical methodology needed to get a restricted forecast with Vector Auto-Regression (VAR) models. In section 2.3 we show how a matrix associated with the uncertainty when imposing unbinding restrictions can be used to cancel out both model dynamics and interactions between restrictions. This fact allows us to obtain the JCT as the sum of several single compatibility tests (SCT), one for each individual restriction. Section 2.4 presents a finite sample Monte Carlo study of the JCT in order to validate its performance. This study led us to an adjustment for the fact that the parameters are estimated. The corresponding restricted forecast formulas for the feasible JCT are also derived. In section 2.5 we illustrate the methodology with an empirical application that uses a model for the Mexican Economy and introduces the economic targets for year 2003, announced by the Mexican Government at the end of 2002. Finally, section 2.6 presents some conclusions.

2.2 Methodology

Many economic time series may tend to move up or down over time in a non-stationary way, but groups of variables may drift together. A multiple time series model is useful to analyze the relationships among these variables and cointegration analysis helps to discover the linear relationships that hold over the long-run. If the variables are cointegrated the basic tool to use for analysis and forecasting is a VEC model. The forecast restricted by extra-model information as well as a compatibility test of unrestricted forecasts and extra-model information are described.

2.2.1 The VEC Model

Let $\mathbf{y}_t = (y_{1t}, \ldots, y_{kt})'$ be a $k \times 1$ vector of random variables at time t and let us assume that $\mathbf{y}_t$ follows a finite pth-order Gaussian VAR model,

$$\mathbf{y}_t = \Lambda \mathbf{D}_t + \Pi_1 \mathbf{y}_{t-1} + \cdots + \Pi_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \tag{2.1}$$

where Π_i is a $k \times k$ coefficient matrix for $i = 1, \ldots, p$, $\mathbf{D}_t = (D_{1t}, \ldots, D_{nt})'$ is an $n \times 1$ vector that includes both deterministic variables to account for seasonality and intervention effects, as well as exogenous variables with respect to $\mathbf{y}_t$. $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{kt})'$ is a k-dimensional independent and identically distributed $N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ random error vector with time-invariant positive-definite covariance matrix $\mathrm{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Sigma}_\varepsilon$. The effects of $\mathbf{D}_t$ on $\mathbf{y}_t$ are captured by the parameter matrix Λ of order $k \times n$.

Model (2.1) can also be written in terms of the lag operator $\mathbf{B}$ as

$$\boldsymbol{\Pi}(\mathbf{B})\,\mathbf{y}_t = \Lambda \mathbf{D}_t + \boldsymbol{\varepsilon}_t \tag{2.2}$$

where $\boldsymbol{\Pi}(\mathbf{B}) = I - \Pi_1 \mathbf{B} - \Pi_2 \mathbf{B}^2 - \cdots - \Pi_p \mathbf{B}^p$ is a matrix polynomial of order p on the lag operator $\mathbf{B}$ such that $\mathbf{B}^n \mathbf{y}_t = \mathbf{y}_{t-n}$ for $n \in \mathbb{N}$. This model can be reparametrized as

$$\boldsymbol{\Pi}^*(\mathbf{B})\,\Delta \mathbf{y}_t = \Lambda \mathbf{D}_t - \boldsymbol{\Pi}(1)\,\mathbf{B}\mathbf{y}_t + \boldsymbol{\varepsilon}_t, \tag{2.3}$$

where Δ is the first difference operator, $\boldsymbol{\Pi}(1) = I - \Pi_1 - \cdots - \Pi_p$ and $\boldsymbol{\Pi}^*(\mathbf{B})$ is a matrix polynomial of order $p-1$ on the lag operator $\mathbf{B}$ such that $\boldsymbol{\Pi}(\mathbf{B}) =$

$\mathbf{\Pi}(1)\,\mathbf{B}+\mathbf{\Pi}^*(\mathbf{B})\,\Delta$. Of course, it is also possible to obtain the VAR in levels from the VAR in differences. Thus equations (2.2) and (2.3) are equivalent representations of the same stochastic process (see for instance Clements and Hendry, 2004, section 8.2).

When the variables in the VAR model are integrated of order $d \geq 1$, written as $I(d)$, Ordinary Least Squares (OLS) estimation of the parameters in model (2.2) is subjected to hazards typical of regressions involving nonstationary variables, see Park and Phillips (1988, 1989). The variables of a k-dimensional process $\mathbf{y}_t$ are cointegrated of order (d, b), briefly $\mathbf{y}_t \sim$CI(d, b), if all components of $\mathbf{y}_t$ are $I(d)$ and there exists a linear combination $\mathbf{cy}_t$, with $\mathbf{c} = (c_1, \ldots, c_k) \neq \mathbf{0}$, which is $I(d-b)$. We assume $\det(\Pi(x)) \neq 0$ for $|x| < 1$, that is, all the zeros are on or outside the unit circle. If the determinantal polynomial has unit roots with multiplicity one, then the variables of $\mathbf{y}_t$ are $I(1)$. Moreover, we assume that the rank of $\Pi(1)$ is r, implying that there are $d = k - r$ unit roots in the system. When $r > 0$ the variables are cointegrated, in the sense that there exists a linear combination $\beta'\mathbf{y}_t$, with $\beta = (\beta_1, \ldots, \beta_k)' \neq \mathbf{0}$, which is $I(0)$. In that case $-\Pi(1) = \alpha\beta'$ where α and β are $k \times r$ full rank matrices. The rows of β' are then referred as the cointegration vectors of the system.

A result due to Engle and Granger (1987), known as Granger's Representation Theorem, is relevant here. It states that if the k variables of $\mathbf{y}_t$ are CI$(1, 1)$ then there exists a VEC representation for the system, that is, equation (2.3) can be written as

$$\Pi^*(\mathbf{B})\,\Delta\mathbf{y}_t = \Lambda\mathbf{D}_t + \alpha\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where $\mathbf{z}_t = \beta'\mathbf{y}_t$ is stationary, the term $\mathbf{\Pi}^*(\mathbf{B})\,\Delta\mathbf{y}_t$ in this model captures the short-run relationships among the variables and $-\mathbf{\Pi}(1)\,\mathbf{B}\mathbf{y}_t = \alpha\mathbf{z}_{t-1}$ captures the long-run relationships.

To introduce the restricted forecasting methodology with VEC models, let us assume first that the model and its parameters are known. Therefore we will not consider such issues as specification, estimation or validation of the model, although they must be considered in practical applications. In section 2.4 we will address the problem of forecasting with estimated parameters. Let us start with the $kT \times 1$ vector $\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_T')'$ that contains all the past information of the multiple time series and let the $kH \times 1$ vector $\mathbf{Y}_F = \left(\mathbf{y}_{T+1}', \ldots, \mathbf{y}_{T+H}'\right)'$ contain the $H \geq 1$ future values to be forecasted for each series. The optimum (in MSE sense) linear forecast of $\mathbf{y}_{T+h}$ for $h = 1, \ldots, H$, is its conditional expectation

$$\mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \Lambda\mathbf{D}_{T+h} + \Pi_1\mathrm{E}(\mathbf{y}_{T+h-1}|\mathbf{Y}) + \cdots + \Pi_p\mathrm{E}(\mathbf{y}_{T+h-p}|\mathbf{Y}) \tag{2.4}$$

where $\mathrm{E}(\mathbf{y}_{T+h-i}|\mathbf{Y}) = \mathbf{y}_{T+h-i}$ for $i \geq h$. Such a forecast produces the forecast error vector

$$\mathbf{y}_{T+h} - \mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \sum_{j=0}^{h-1}\Psi_j\boldsymbol{\varepsilon}_{T+h-j}, \text{ for } h = 1, \ldots, H \tag{2.5}$$

where $\Psi_j = \sum_{k=1}^{j}\Psi_{j-k}\Pi_k$ with $\Psi_0 = I$ and $\Pi_k = 0$ for $k > p$. It should be stressed that, by virtue of the equivalence between (2.2) and (2.3), the forecast provided by (2.4) make use of all the information in the VEC model.

Expression (2.5) can be written as

$$\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \mathbf{\Psi}\boldsymbol{\epsilon}_F \tag{2.6}$$

where $\boldsymbol{\epsilon}_F = \left(\boldsymbol{\varepsilon}'_{T+1}, \ldots, \boldsymbol{\varepsilon}'_{T+H}\right)' \sim N\left(\mathbf{0}, I_H \otimes \mathbf{\Sigma}_\varepsilon\right)$ is a $kH \times 1$ vector, with $\otimes$ the Kronecker product. The $kH \times kH$ matrix $\mathbf{\Psi}$ is lower triangular with I_k in its main diagonal, Ψ_1 in its first subdiagonal, Ψ_2 in the second subdiagonal and so on. Thus the optimum unrestricted forecast of $\mathbf{Y}_F$ is its conditional expectation $\mathbf{Y}_{F,H}$ =E($\mathbf{Y}_F|\mathbf{Y}$), whose MSE is given by

$$\begin{aligned} \Sigma_{\mathbf{Y_{F,H}}} &\equiv \mathrm{E}[(\mathbf{Y}_F - \mathbf{Y}_{F,H})(\mathbf{Y}_F - \mathbf{Y}_{F,H})'|\mathbf{Y}] \\ &= \mathbf{\Psi}\left(I_H \otimes \mathbf{\Sigma}_\varepsilon\right)\mathbf{\Psi}'. \end{aligned}$$

The forecast MSE for an integrated process is generally unbounded as the horizon H goes to infinity, that is, the forecast uncertainty increases without limit, see for example Lütkepohl (1991, section 11.3).

2.2.2 Stochastic linear restrictions on future values

Suppose we have additional information on the future of $\mathbf{y}_t$. For example, when the government wants to reach some economic targets, we assume the forecast of the variables $\mathbf{y}_t$ are restricted by those targets, at least to produce a scenario. We also assume that the additional information is in the form of a stochastic linear restriction

$$\mathbf{R} = \mathbf{C}\mathbf{Y}_F + \mathbf{u}. \tag{2.7}$$

Here $\mathbf{C}$ is an $M \times kH$ matrix of full row rank, $\mathbf{R} = (r_1, \ldots, r_M)'$ is the $M \times 1$ vector of values that the linear combinations take on and $\mathbf{u} = (u_1, \ldots, u_M)' \sim N(\mathbf{0}, \Sigma_{\mathbf{u}})$ is an $M \times 1$ random vector, with the ijth element of $\Sigma_{\mathbf{u}}$ given by $\sigma_{ij,u} = \mathrm{cov}(u_i, u_j)$, for $i, j = 1, \ldots, M$.

To illustrate this situation let us consider a bivariate VAR system where the first variable must satisfy an isolated stochastic linear restriction as $r = y_{1,T+H} + u$, with $u \sim N(0, \sigma_u)$. Then in terms of (2.7) we have $\mathbf{R} = r$, $\mathbf{u} = u$ and $\mathbf{Y}_F = (y_{1,T+1}, y_{2,T+1}, \ldots, y_{1,T+H-1}, y_{2,T+H-1}, y_{1,T+H}, y_{2,T+H})'$ with $\mathbf{C} = (0, 0, \ldots, 0, 0, 1, 0)$ where there are first $H-1$ pairs of zeros and then the pair $1, 0$. This kind of restriction arises when the government announces a target for the first variable for the end of the year.

Another example considers the difference between two future values of the first variable $r_1 = y_{1,T+H} - y_{1,T+1} + u_1$ with $u_1 \sim N(0, \sigma_{11,u})$, while the second variable has a restriction on the average of its future values $r_2 = \sum y_{2,T+i}/H + u_2$ with $u_2 \sim N(0, \sigma_{22,u})$. Further, if u_1 and u_2 are uncorrelated, these two restrictions can be written as

$$\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \mathbf{C}\mathbf{Y}_F + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}; \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11,u} & 0 \\ 0 & \sigma_{22,u} \end{pmatrix} \right],$$

where

$$\mathbf{C} = \begin{pmatrix} -1 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & \frac{1}{H} & \cdots & 0 & \frac{1}{H} & 0 & \frac{1}{H} \end{pmatrix}.$$

In the empirical example shown in section 2.5 we face the problem of imposing three restrictions on the future values of a six-variable system of the Mexican economy. One of those restrictions is of the difference type and the other two involve averages.

2.2.3 Restricted forecasts

When forecasting $\mathbf{Y}_F$ we should consider both historical information, $\mathbf{Y}$, and additional information, $\mathbf{R}$. The optimum forecast that takes into account both sources of information is $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y},\mathbf{R})$ which will be denoted as $\mathbf{Y}^R_{F,H}$. It can be shown, see for instance Nieto and Guerrero's (1995) Theorem 1, that

$$\mathbf{Y}^R_{F,H} = \mathbf{Y}_{F,H} + \mathbf{A}(\mathbf{R} - \mathbf{C}\mathbf{Y}_{F,H})$$

where $\mathbf{A} = \Sigma_{\mathbf{Y}_{F,H}}\mathbf{C}'\mathbf{\Omega}^{-1}$and $\mathbf{\Omega} \equiv \mathbf{C}\Sigma_{\mathbf{Y}_{F,H}}\mathbf{C}'+\Sigma_{\mathbf{u}}$ is an $M \times M$ symmetric positive-semidefinite matrix. Moreover, its MSE is given by

$$\begin{aligned}\mathrm{MSE}\left(\mathbf{Y}^R_{F,H}\right) &\equiv \mathrm{E}\left[\left(\mathbf{Y}_F - \mathbf{Y}^R_{F,H}\right)\left(\mathbf{Y}_F - \mathbf{Y}^R_{F,H}\right)'\middle|\, \mathbf{Y},\mathbf{R}\right] \\ &= \left(I - \mathbf{AC}\right)\Sigma_{\mathbf{Y}_{F,H}}.\end{aligned}$$

Rearranging terms we have

$$\Sigma_{\mathbf{Y}_{F,H}} = \mathrm{MSE}\left(\mathbf{Y}^R_{F,H}\right) + \mathbf{AC}\Sigma_{\mathbf{Y}_{F,H}},$$

since $\mathrm{MSE}\left(\mathbf{Y}^R_{F,H}\right)$ and $\mathbf{AC}\Sigma_{\mathbf{Y}_{F,H}}$ are positive semidefinite matrices, we see that $\mathbf{Y}^R_{F,H}$ is at least as precise as $\mathbf{Y}_{F,H}$.

2.2.4 Compatibility test

In spite of the optimality of $\mathbf{Y}_{F,H}^{R}$, we should test if the two sources of information are compatible with each other, in which case the combination makes sense. Guerrero *et al.* (2005) proposed to define an $M \times 1$ distance vector

$$\mathbf{d} \equiv \mathbf{R} - \mathbf{C}\mathbf{Y}_{F,H} = \mathbf{C}\boldsymbol{\Psi}\boldsymbol{\epsilon}_F + \mathbf{u}, \tag{2.8}$$

whose elements are denoted by d_i, $i = 1, \ldots, M$. Then $\mathbf{R}$ and $\mathbf{Y}_{F,H}$ are said to be compatible if the distance vector is close to zero. From the normality assumption of $\boldsymbol{\epsilon}_F$ and $\mathbf{u}$, we know that before observing the values of $\mathbf{Y}_{F,H}$ and $\mathbf{R}$,

$$\mathbf{d} \sim N\left(\mathbf{0}, \boldsymbol{\Omega}\right).$$

Thus the following JCT statistic arises naturally

$$\mathbf{K} \equiv \mathbf{d}'\boldsymbol{\Omega}^{-1}\mathbf{d} \sim \chi^2_M. \tag{2.9}$$

Then, after observing the values of $\mathbf{Y}_{F,H}$ we say that $\mathbf{R}$ is in the compatibility region at level α if

$$\mathbf{K}_{calc} = \left(\mathbf{R} - \mathbf{C}\mathbf{Y}_{F,H}\right)' \boldsymbol{\Omega}^{-1} \left(\mathbf{R} - \mathbf{C}\mathbf{Y}_{F,H}\right) \leq \chi^2_M(\alpha),$$

with $\chi^2_M(\alpha)$ the $(1-\alpha)$th quantile of the χ^2_M distribution. This distribution is exact when the matrices $\boldsymbol{\Psi}$, $\boldsymbol{\Sigma}_\varepsilon$ and $\Sigma_{\mathbf{u}}$ are known. When they are consistently estimated, the χ^2 distribution will be valid asymptotically. It should be noticed that even though the $\mathbf{y}_t$ variables are $I(1)$, the statistic $\mathbf{K}$ follows a standard χ^2

distribution since it is derived from the distance vector $\mathbf{d}$, which has a normal distribution because of (2.6) and (2.7).

Let us note that the JCT is an omnibus test that does not consider any specific alternative hypothesis. If the JCT leads us to conclude incompatibility between sources of information, one reason could be that the additional information imposed a binding restriction ($\Sigma_{\mathbf{u}} = \mathbf{0}$), in which case the possibility of having $\Sigma_{\mathbf{u}} \neq \mathbf{0}$ arises as an alternative. It can also happen that only some of the individual restrictions are incompatible with their corresponding unrestricted forecasts. In such a case we require individual compatibility tests that enable us to appreciate the contribution of each individual restriction to the JCT. These last two issues will be studied in the following sections.

2.3 Additivity of the compatibility test

In what follows we shall refer to $\mathbf{K}$ as the JCT and the test of a single restriction will be called an SCT. In this section we analyze how the M SCTs are related to $\mathbf{K}$. Then we propose to modify the matrix $\Sigma_{\mathbf{u}}$ to achieve additivity.

2.3.1 Decomposition of the JCT into single tests

In analogy with (2.8), for the ith restriction, with $i = 1, \ldots, M$, we get

$$\begin{aligned} d_i &\equiv r_i - C_i \mathbf{Y}_{F,H} \\ &= C_i \boldsymbol{\Psi} \boldsymbol{\epsilon}_F + u_i, \end{aligned}$$

where C_i is the ith row of $\mathbf{C}$ (representing the linear combinations of $\mathbf{Y}_F$ involved in this single restriction) and $u_i \sim N(0, \sigma_{ii,u})$. From here we have that $d_i \sim N(0, \omega_{ii})$, where $\omega_{ii} \equiv C_i \Sigma_{\mathbf{Y}_{F,H}} C_i' + \sigma_{ii,u}$ is the iith element of $\boldsymbol{\Omega}$. As before, we define the SCT of the ith restriction as

$$K_{ii} = d_i \omega_{ii}^{-1} d_i = \omega_{ii}^{-1} d_i^2 \sim \chi_1^2.$$

The following proposition states that, in general, the JCT cannot be obtained simply as the sum of the SCTs.

Proposition 1 *Let $\Sigma_{\mathbf{Y}_{F,H}}$ be a symmetric real positive-definite matrix and $\mathbf{C}$ be of full row rank. Then, the JCT can be written in terms of the SCTs as*

$$\mathbf{K} = \sum_{i=1}^{M} \overline{\omega}_{ii} \omega_{ii} K_{ii} + 2 \sum_{i<j}^{M} \overline{\omega}_{ij} \omega_{ij} K_{ij}, \tag{2.10}$$

where ω_{ij} and $\overline{\omega}_{ij}$ denote the $(i,j)-th$ elements of $\boldsymbol{\Omega}$ and $\boldsymbol{\Omega}^{-1}$ respectively. Besides $K_{ij} \equiv \omega_{ij}^{-1} d_i d_j$ if $\omega_{ij} \neq 0$ and $K_{ij} \equiv 0$ if $\omega_{ij} = 0$, for $i, j = 1, \ldots, M$.

Proof: Since $\text{rank}(\mathbf{C}) = M$, $\mathbf{C} \Sigma_{\mathbf{Y}_{F,H}} \mathbf{C}'$ is a symmetric positive-definite matrix. Now, $\Sigma_{\mathbf{u}}$ is the symmetric and positive-semidefinite covariance matrix of $\mathbf{u}$, thus $\boldsymbol{\Omega} = \mathbf{C} \Sigma_{\mathbf{Y}_{F,H}} \mathbf{C}' + \Sigma_{\mathbf{u}}$ is also a symmetric positive-definite matrix. Which means that det($\mathbf{\Omega}$) >0, so that $\mathbf{\Omega}^{-1}$ exists. Hence, substituting the elements of $\mathbf{d}$ and $\mathbf{\Omega}^{-1}$ directly into (2.9) the JCT takes the indicated form. □

When the matrix $\mathbf{\Omega}$ involved in the distance vector $\mathbf{d}$ is diagonal, we get

$$\begin{aligned}\mathbf{K} &= \omega_{11}^{-1}d_1^2 + \cdots + \omega_{MM}^{-1}d_M^2 \\ &= K_{11} + K_{22} + \cdots + K_{MM}\end{aligned}$$

that is, the SCTs add up to JCT.

The converse is not true, that is, $\mathbf{K} = K_{11} + \cdots + K_{MM}$ does not imply that $\mathbf{\Omega}$ is diagonal. To see this let us consider a simple case with $M = 2$, so that $\mathbf{K} = K_{11} + K_{22} = (\omega_{22}d_1^2 + \omega_{11}d_2^2) / (\omega_{11}\omega_{22})$. Now, from (2.10) the general form of JCT is $\mathbf{K} = (\omega_{22}d_1^2 + \omega_{11}d_2^2 - 2\omega_{21}d_1d_2) / (\omega_{11}\omega_{22} - 2\omega_{21})$. Thus, solving for the correlation term $\omega_{21} = \omega_{12}$, the JCT will be the sum of the SCTs only if $\omega_{21} = 0$ or $\omega_{21} = 2d_1d_2/(K_{11} + K_{22})$. Therefore, this example shows that there exists a non-diagonal matrix $\mathbf{\Omega}$ that yields $\mathbf{K} = K_{11} + K_{22}$.

2.3.2 Decomposition of $\mathbf{K}$

It should be clear that many factors that affect the economic system are not under control of the government and usually imply some uncertainty on the targets. The problem in our case is how to assign this uncertainty. One possibility is that $\Sigma_{\mathbf{u}}$ be given by the same external source that provides the linear restrictions, for instance, when it comes from a competing forecasting method. Otherwise, choosing this matrix should be done with care in order not to lose precision in

$\mathbf{Y}^R_{F,H}$. For example, Guerrero and Peña (2003) considered that this matrix has a diagonal form and chose its elements to make the restrictions compatible with the historical data. Here we propose another alternative.

We suggest to modify the matrix $\Sigma_{\mathbf{u}}$ to cancel out model dynamics and interactions between restrictions and thus obtain additivity of the SCTs with respect to the JCT. To that end, let us define

$$\Sigma_{\mathbf{u}}(a) \equiv \left[a\mathrm{diag}(Q) + \mathrm{diag}\left(\mathbf{C}\Sigma_{\mathbf{Y}_{F,H}}\mathbf{C}'\right) - \mathbf{C}\Sigma_{\mathbf{Y}_{F,H}}\mathbf{C}'\right], \tag{2.11}$$

for some real parameter $a > 0$ and Q an $M \times 1$ vector whose elements are denoted by $q_i > 0$ for $i = 1, \ldots, M$.[1] These q_i values will be interpreted as weights associated to the uncertainty of the restrictions involved, as it will be seen in the following section. More precisely, matrix (2.11) takes the form

$$\Sigma_{\mathbf{u}}(a) = \begin{pmatrix} aq_1 & -\omega_{12} & \ldots & -\omega_{1M} \\ \ldots & \ldots & \ldots & \ldots \\ -\omega_{M1} & -\omega_{M2} & \ldots & aq_M \end{pmatrix}.$$

Let $\lambda_1(a) \geq \lambda_2(a) \geq \cdots \geq \lambda_M(a)$ denote the eigenvalues of this matrix and define the modified matrix as

$$\Sigma^*_{\mathbf{u}} \equiv \Sigma_{\mathbf{u}}(a^*) \quad \text{s.t.} \quad \lambda_M(a^*) = 0 \tag{2.12}$$

[1] If $A = (a_{ij})$ is an $n \times n$ matrix and $\mathbf{a} = (a_i)$ is an $n \times 1$ vector, the operators $\mathrm{diag}(A) = \begin{pmatrix} a_{11} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & a_{nn} \end{pmatrix}$ and $\mathrm{diag}(\mathbf{a}) = \begin{pmatrix} a_1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & a_n \end{pmatrix}$ produce $n \times n$ diagonal matrices.

which, by construction, is a symmetric positive semidefinite matrix and defines a covariance matrix. The function $\lambda_M(a)$ may have more than one zero, in which case the solution of (2.12) will not be unique. Below we shall see that this situation is avoided by the form of $\Sigma_{\mathbf{u}}(a)$. In fact, this follows from the linear dependence of the eigenvalues on the parameter a.

Let $\mathbf{v}_i$ be an eigenvector of $\Sigma_{\mathbf{u}}$ corresponding to λ_i, for $i = 1, \ldots, M$, thus

$$\Sigma_{\mathbf{u}}\mathbf{v}_i = \lambda_i \mathbf{v}_i; \quad \text{with} \quad \mathbf{v}_i'\mathbf{v}_i = 1.$$

The differential of this expression is

$$\begin{aligned} d(\Sigma_{\mathbf{u}}\mathbf{v}_i) &= \operatorname{diag}(Q)\, d(a)\, \mathbf{v}_i + \Sigma_{\mathbf{u}} d(\mathbf{v}_i) \\ &= d(\lambda_i)\, \mathbf{v}_i + \lambda_i d(\mathbf{v}_i) = d(\lambda_i \mathbf{v}_i) \end{aligned}$$

see Magnus and Neudecker (2002). Premultiplying by $\mathbf{v}_i'$ we have

$$\mathbf{v}_i' \operatorname{diag}(Q)\, d(a)\, \mathbf{v}_i + \mathbf{v}_i'\Sigma_{\mathbf{u}} d(\mathbf{v}_i) = \mathbf{v}_i' d(\lambda_i)\, \mathbf{v}_i + \mathbf{v}_i'\lambda_i d(\mathbf{v}_i).$$

Hence, since $\Sigma_{\mathbf{u}}$ is symmetric we get

$$\mathbf{v}_i' \operatorname{diag}(Q)\, d(a)\, \mathbf{v}_i = \mathbf{v}_i' d(\lambda_i)\, \mathbf{v}_i,$$

and

$$d(\lambda_i) = [\mathbf{v}_i' \operatorname{diag}(Q)\, \mathbf{v}_i]\, d(a),$$

because the eigenvector $\mathbf{v}_i$ is normalized by $\mathbf{v}_i'\mathbf{v}_i = 1$ and a and λ_i are scalars.

The last expression implies that

$$\frac{d\lambda_i(a)}{da} = c \quad \text{for } i = 1, \ldots, M, \tag{2.13}$$

Remark 1 *It should be noticed that additivity is achieved by introducing uncertainty in the restrictions. Although we introduce the minimum amount of uncertainty in the sense of (2.12) it could be that an incompatible binding restriction becomes a compatible unbinding restriction. This fact may be considered a drawback of this proposal, nevertheless it may be useful in the practical applications as can be seen in the empirical application of this paper.*

2.4 A feasible JCT for estimated processes

We study the finite sample properties of the JCT by way of a numerical simulation of a bivariate system. Three typical restrictions and four sample sizes $T = 20, 50, 100$ and 200 are used. This leads us to consider first that the process is estimated and second that the covariance matrix of the distance vector should take into account the sample size. Thus a feasible JCT for estimated processes is obtained and its finite sample properties are studied in a similar way. Another simulation with a model used in the following section for the Mexican Economy was carried out for the JCT and for the feasible JCT in order to validate the empirical application of these tests. We did not carry out a simulation study for the univariate situation because Box and Tiao (1976) did that for a statistic similar in nature to $\mathbf{K}$. In fact, they found that the distribution of their statistic was Chi-square as in (2.9). However, for an estimated AR(p) model with n observations the "estimation errors inflate the mean value of χ^2... by a factor approximating

$1 + (p/n)$."

2.4.1 A Monte Carlo study of the JCT

We assume that the process follows the bivariate VAR(2) model given in Lütkepohl (1991, 3.2.25-3.2.26).

$$\mathbf{y}_t = \begin{pmatrix} 0.02 \\ 0.03 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \tag{2.14}$$

with error covariance matrix

$$\Sigma_\varepsilon = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} \times 10^{-4}. \tag{2.15}$$

As is usual in Monte Carlo experiments, we lose in generality in order to gain in understanding the small sample properties of a particular case. Suppose that the government announces at the end of the year the economic targets for next year and that data are quarterly (data for the current quarter are still unavailable). We consider three typical linear combinations of $\mathbf{Y}_F$ for the stochastic restriction

$$\mathbf{C}_1 = \begin{pmatrix} 0 & 0 & ... & 0 & 0 & 1 & 0 \\ 0 & 0 & ... & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{C}_2 = \begin{pmatrix} 0 & 0 & \frac{1}{4} & 0 & ... & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} & ... & 0 & \frac{1}{4} \end{pmatrix} \tag{2.16}$$

and

$$\mathbf{C}_3 = \begin{pmatrix} -1 & 0 & 0 & 0 & ... & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & ... & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.17}$$

The $\mathbf{C}_1$ matrix puts a restriction at the end of the forecast horizon. This is useful when the government announces the value of a variable for the end of the following year. $\mathbf{C}_2$ applies when the government announces the average value for a variable in the following year. If the 1/4 values are replaced by 1 the restriction is on the sum of the variable. This specification will be used in the empirical illustration when dealing with the inflation rate (in logs). Finally, $\mathbf{C}_3$ restricts the difference of values for the end of the year. This is used when dealing with the rate of growth of a variable. We apply this restriction to GDP (in logs) in the empirical illustration.

The Monte Carlo experiment was done with the following algorithm and program routines were written in Matlab 6.5-Release 13 (MathWorks, Inc. Software). Given $\mathbf{C}$, $\Sigma_{\mathbf{u}}$, T, H, and p:

1. *Generate a series* $\{\mathbf{y}_t\}_{t=-p+1}^{T+H}$ *with* (2.14)-(2.15) *and initial values* $E(\mathbf{y}_t) = (0.07027, 0.15135)'$. *Following the partition used by Lütkepohl we have a time series* $\mathbf{y}_1, \ldots, \mathbf{y}_T$ *of length* T *as well as a presample* $\mathbf{y}_{-p+1}, \ldots, \mathbf{y}_0$. *The future values of the series* $\mathbf{Y}_F$ *are given by* $\mathbf{y}_T, \ldots, \mathbf{y}_{T+H}$.

2. *Estimate the* VAR(p) *for the process. Compute* $\mathbf{Y}_{F,H}$ and $\Sigma_{\mathbf{Y}_{F,H}}$.

3. *Generate a random vector* $\mathbf{u} \sim N(\mathbf{0}, \Sigma_{\mathbf{u}})$. *For binding restrictions* $(\Sigma_{\mathbf{u}} = \mathbf{0})$ *take* $\mathbf{u} = \mathbf{0}$.

4. *Compute* $\mathbf{R} = \mathbf{C}\mathbf{Y}_F + \mathbf{u}$.

5. *Compute* $\mathbf{\Omega}$, $\mathbf{d}$ *and* $\mathbf{K}$.

Steps 1-5 were replicated $N = 1000$ times for $p = 2$, $T = 20, 50, 100, 200$, with three restrictions of the form (2.16)-(2.17) and considering $H = 5$ with $\Sigma_{\mathbf{u}} = \mathbf{0}$ fixed. Since $\mathbf{K}$ is distributed as χ_2^2 we know that $P(\mathbf{K} \geq \chi_2^2(\alpha)) = \alpha$. In practice we should expect the proportion of sample values of $\mathbf{K}$ that exceeds the $\chi_2^2(\alpha)$ quantile to be close to α. We denote this proportion by p_α. Table 2.1 reports these values for α= .10, .05 and .01. The p_α values for $T = 20, 50$ and 100 are higher than α. Thus the χ_2^2 distribution leads to over-reject the compatibility hypothesis. For example, when $T = 20$, $\mathbf{C} = \mathbf{C}_1$ and $\alpha = 0.10$ ($\Sigma_{\mathbf{u}} = \mathbf{0}$, $H = 5$). The reported value says that 32% of the trials produced $\mathbf{K}$ statistics above $\chi_2^2(.10)$. Notice that for $T = 20$ and 50, the sampling distribution of $\mathbf{K}$ depends on the restriction imposed. For example, the p_α values with $\mathbf{C}_2$ are higher than those for $\mathbf{C}_1$ and $\mathbf{C}_3$, this may be due to the fact that using the $\mathbf{C}_2$ matrix involves restricting more values of $\mathbf{Y}_F$ than with $\mathbf{C}_1$ or $\mathbf{C}_3$.

This experiment was repeated with the matrix $\Sigma_{\mathbf{u}}^*$ in place of $\Sigma_{\mathbf{u}}$. Since this matrix is computed once the VAR has been estimated, the algorithm described above must be changed. Given $\mathbf{C}$, T, H and p step 3 should be replaced by

3'. *Compute Q as a diagonal matrix, with the squared values of $v = \mathbf{C}(\mathbf{Y}_F - \mathbf{Y}_{F,H})$ in its main diagonal. Compute $\Sigma_{\mathbf{u}}^*$ as defined in* (2.12). *Generate a random vector $\mathbf{u} \sim N(\mathbf{0}, \Sigma_{\mathbf{u}}^*)$.*

The Q matrix computed this way assigns uncertainty to each restriction in accordance to the distance between the linear restriction without uncertainty $\mathbf{C}\mathbf{Y}_F$

Table 2.1: p_α values for the compatibility test with binding restrictions

Sample size	Combination matrix ($H = 5$)								
	$\mathbf{C}_1$			$\mathbf{C}_2$			$\mathbf{C}_3$		
	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$
20	.32	.25	.14	.48	.40	.26	.30	.21	.11
50	.18	.12	.05	.22	.14	.05	.17	.10	.04
100	.14	.09	.02	.14	.09	.03	.14	.09	.02
200	.10	.06	.02	.10	.05	.01	.10	.06	.01

Proportion of $\mathbf{K}_{\text{calc}}$ values exceeding the $\chi^2_2(\alpha)$ quantile with $\Sigma_{\mathbf{u}} = \mathbf{0}$ obtained with $N = 1000$ Monte Carlo replications of (2.14)–(2.15).

(see 2.7) and the linear combinations of the unrestricted forecast $\mathbf{CY}_{F,H}$. Table 2.2 reports the p_α proportions for this experiment.

In summary, the p_α proportions are higher than α for small samples. This result does not depend on whether the restriction is binding or unbinding, but assigning uncertainty in this way leads the p_α values closer to α than those reported for the binding case. In the following section we propose an adjustment for the JCT that also takes into account that the model parameters are estimated. As we will see in the next subsection, this yields an additional improvement for small samples.

Table 2.2: p_α values for the compatibility test with unbinding restrictions

Sample size	Combination matrix ($H=5$)								
	$\mathbf{C}_1$			$\mathbf{C}_2$			$\mathbf{C}_3$		
	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$
20	.27	.19	.08	.35	.28	.17	.23	.17	.08
50	.14	.08	.03	.15	.11	.04	.12	.06	.02
100	.10	.05	.01	.12	.06	.01	.11	.06	.01
200	.10	.04	.01	.10	.06	.02	.10	.05	.01

Proportion of $\mathbf{K}_{\text{calc}}$ values exceeding the $\chi^2_2(\alpha)$ quantile with $\Sigma^*_{\mathbf{u}}$ obtained with $N=1000$ Monte Carlo replications of (2.14)–(2.15).

2.4.2 The VAR forecast for estimated processes

The optimal h-step forecast of (2.1) is given by (2.4), however, if the true parameters $\Theta = (\Lambda, \Pi_1, \ldots, \Pi_p)$ are replaced by their estimators $\widehat{\Theta} = \left(\widehat{\Lambda}, \widehat{\Pi}_1, \ldots, \widehat{\Pi}_p\right)$ we get the forecast

$$\widehat{\mathrm{E}}(\mathbf{y}_{T+h}|\mathbf{Y}) = \widehat{\Lambda}\mathbf{D}_{T+h} + \widehat{\Pi}_1\widehat{\mathrm{E}}(\mathbf{y}_{T+h-1}|\mathbf{Y}) + \cdots + \widehat{\Pi}_p\widehat{\mathrm{E}}(\mathbf{y}_{T+h-p}|\mathbf{Y}) \quad \text{for } h = 1, \ldots, H,$$

where $\widehat{\mathrm{E}}(\mathbf{y}_{T+h-i}|\mathbf{Y}) = \mathbf{y}_{T+h-i}$ for $i \geq h$. By calling $\widehat{\mathbf{Y}}_{F,H}$ the $\widehat{\mathrm{E}}(\mathbf{Y}_F|\mathbf{Y})$ vector we have

$$\begin{aligned} \mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H} &= (\mathbf{Y}_F - \mathbf{Y}_{F,H}) + \left(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\right), \\ &= \boldsymbol{\Psi}\boldsymbol{\epsilon}_F + \left(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\right). \end{aligned} \tag{2.18}$$

Under general conditions for the process $\mathbf{y}_t$, Dufour (1985) proved that the forecast errors have zero conditional mean, $\mathrm{E}\Big(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}|\mathbf{Y}\Big) = \mathbf{0}$, so the forecast is conditionally unbiased even if the parameters are estimated. On the right side of (2.18), all the $\boldsymbol{\varepsilon}_t$ contained in $\boldsymbol{\epsilon}_F$ correspond to periods $t > T$ whereas all the $\mathbf{y}_t$ contained in the second term correspond to $t \leq T$. Therefore the two terms are uncorrelated. Thus the MSE of $\widehat{\mathbf{Y}}_{F,H}$ becomes

$$\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,H}\right) \equiv \Sigma_{\widehat{\mathbf{Y}}_{F,H}} = \Sigma_{\mathbf{Y}_{F,H}} + \mathrm{MSE}\left(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\right), \tag{2.19}$$

To evaluate the second term on the right hand side of this equation we need the distribution of $\widehat{\Theta}$. Since small sample distributions of VAR estimators are not available, we cannot hope to get more than an asymptotic distribution for $\mathrm{MSE}\Big(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\Big)$. To obtain this we proceed as in Lütkepohl (1991, section 3.5). So, let $\boldsymbol{\beta} \equiv \mathrm{vec}(\Theta)$ and $\widehat{\boldsymbol{\beta}} \equiv \mathrm{vec}(\widehat{\Theta})$ be its OLS estimator, whose asymptotic covariance matrix is $\Sigma_{\widehat{\boldsymbol{\beta}}}$ and

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \xrightarrow{d} N\left(\mathbf{0}, \Sigma_{\widehat{\boldsymbol{\beta}}}\right)$$

Then, under quite general conditions we have that, conditional on $\mathbf{Y}$,

$$\sqrt{T}\left(\widehat{\mathbf{Y}}_{F,H} - \mathbf{Y}_{F,H}|\mathbf{Y}\right) \xrightarrow{d} N\left(\mathbf{0}, \frac{\partial \mathbf{Y}_{F,H}}{\partial \boldsymbol{\beta}'}\Sigma_{\widehat{\boldsymbol{\beta}}}\frac{\partial \mathbf{Y}'_{F,H}}{\partial \boldsymbol{\beta}}\right). \tag{2.20}$$

From here, an approximation to the $\mathrm{MSE}\Big(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\Big)$ is $T^{-1}\widetilde{\Sigma}$, where

$$\widetilde{\Sigma} = \mathrm{E}\left(\frac{\partial \mathbf{Y}_{F,H}}{\partial \boldsymbol{\beta}'}\Sigma_{\widehat{\boldsymbol{\beta}}}\frac{\partial \mathbf{Y}'_{F,H}}{\partial \boldsymbol{\beta}}\right). \tag{2.21}$$

Therefore (2.19) becomes

$$\Sigma_{\widehat{\mathbf{Y}}_{F,H}} \approx \Sigma_{\mathbf{Y}_{F,H}} + T^{-1}\widetilde{\Sigma},$$

the explicit expression for $\widetilde{\Sigma}$ is derived in Appendix A.

2.4.2.1 Restricted forecast with an estimated process

The optimum restricted forecast with an estimated process is

$$\widehat{\mathbf{Y}}_{F,H}^{R} = \widehat{\mathbf{Y}}_{F,H} + \widehat{\mathbf{A}}\left[\mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right]$$

where $\widehat{\mathbf{A}} = \Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}'\widehat{\mathbf{\Omega}}^{-1}$and $\widehat{\mathbf{\Omega}} \equiv \mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}' + \Sigma_{\mathbf{u}}$. Its MSE is given by

$$\begin{aligned}
\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,H}^{R}\right) &\equiv \mathrm{E}\left[\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}^{R}\right)\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}^{R}\right)'\middle|\,\mathbf{Y},\mathbf{R}\right] \\
&= \left(I - \widehat{\mathbf{A}}\mathbf{C}\right)\Sigma_{\widehat{\mathbf{Y}}_{F,H}}.
\end{aligned}$$

As before, we have that $\widehat{\mathbf{Y}}_{F,H}^{R}$ is at least as precise as $\widehat{\mathbf{Y}}_{F,H}$. These expressions are derived in Appendix B.

2.4.2.2 The feasible JCT

A compatibility test must consider the distribution of the distance vector of an estimated process as well, that is

$$\begin{aligned}
\widehat{\mathbf{d}} &\equiv \mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H} \\
&= \mathbf{C}\boldsymbol{\Psi}\boldsymbol{\epsilon}_F + \mathbf{C}\left[\sqrt{T}\left(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\right)\right]/\sqrt{T} + \mathbf{u}.
\end{aligned} \tag{2.22}$$

From the normality assumption of $\boldsymbol{\epsilon}_F$, $\mathbf{u}$ and (2.20) we have, approximately

$$\widehat{\mathbf{d}} \sim N\left(\mathbf{0}, \mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}' + \Sigma_{\mathbf{u}}\right).$$

Thus, for estimated processes the statistic $\widehat{\mathbf{d}}'\left(\mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}' + \Sigma_{\mathbf{u}}\right)^{-1}\widehat{\mathbf{d}}$ is plausible. Consistent estimators of the MSE matrices are obtained by replacing the unknown parameters by their estimators. The resulting estimator of $\Sigma_{\widehat{\mathbf{Y}}_{F,H}}$ will be denoted by $\widehat{\Sigma}_{\widehat{\mathbf{Y}}_{F,H}}$.

In addition to the estimated process adjustment we should adjust the statistic for using an estimated rather than a known covariance matrix. This leads us to divide the statistic by its degrees of freedom and then using an F distribution to make inferences. Finally, the feasible JCT gets defined as

$$\overline{\mathbf{K}} = \widehat{\mathbf{d}}'\overline{\boldsymbol{\Omega}}^{-1}\widehat{\mathbf{d}}/M \sim F_{M,T-Mp-1}$$

where $\overline{\boldsymbol{\Omega}} \equiv \mathbf{C}\widehat{\Sigma}_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}' + \Sigma_{\mathbf{u}}$.

Thus, $\mathbf{R}$ is not in the compatibility region at level of significance α if

$$\overline{\mathbf{K}}_{calc} \geqslant F_{M,T-Mp-1}(\alpha)$$

with $F_{M,T-Mp-1}(\alpha)$ the $(1-\alpha)$th quantile of the $F_{M,T-Mp-1}$ distribution. Since $M\overline{\mathbf{K}}$ has the same quadratic form as $\mathbf{K}$, the results obtained in Section 2.3 remain valid. For example, the modified matrix (2.12) makes use of

$$\Sigma_{\mathbf{u}}(a) \equiv \left[a\mathrm{diag}(Q) + \mathrm{diag}\left(\mathbf{C}\widehat{\Sigma}_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}'\right) - \mathbf{C}\widehat{\Sigma}_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}'\right]$$

and the additivity property of the feasible JCT yields

$$M\overline{\mathbf{K}} = \overline{K}_{11} + \overline{K}_{22} + \cdots + \overline{K}_{MM}.$$

2.4.3 A Monte Carlo study for the feasible JCT

Table 2.3 presents the proportion, p_α, of sample values of $\overline{\mathbf{K}}$ exceeding the quantile $F_{M,T-Mp-1}(\alpha)$. The p_α values are closer to α than those reported for the JCT in Table 1. For sample sizes $T \geq 50$ it can be seen that $F_{M,T-Mp-1}$ approximates well the distribution of $\overline{\mathbf{K}}$. As before, it is also clear that the sampling distribution of $\overline{\mathbf{K}}$ depends on the restriction used.

The Monte Carlo experiment with $\overline{\mathbf{K}}$ was also done for the unbinding case with $\Sigma^*_{\mathbf{u}}$. Table 2.4 reports the corresponding proportions for this case. Magnitudes of the p_α values are similar to those in Table 2.3. So, the $F_{M,T-Mp-1}$ distribution can also be considered a good approximation for the finite distribution of $\overline{\mathbf{K}}$ for sample sizes $T \geq 50$, but its sampling distribution depends on the type of restriction employed.

Remark 2 *The simulations provide some evidence that the feasible JCT sta-*

Table 2.3: p_α values for the feasible compatibility test with binding restrictions

Sample size	Combination matrix ($H = 5$)								
	$\mathbf{C}_1$			$\mathbf{C}_2$			$\mathbf{C}_3$		
	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$
20	.23	.14	.05	.30	.24	.13	.18	.11	.04
50	.13	.07	.02	.17	.11	.03	.12	.07	.01
100	.12	.06	.02	.15	.09	.03	.11	.06	.01
200	.11	.07	.01	.12	.06	.01	.11	.05	.01

Proportion of $\overline{\mathbf{K}}_{\text{calc}}$ values exceeding the $F_{M,T-Mp-1}(\alpha)$ quantile with $\Sigma_{\mathbf{u}} = \mathbf{0}$ obtained with $N = 1000$ Monte Carlo replications of (2.14)–(2.15).

Table 2.4: p_α values for the feasible compatibility test with unbinding restrictions

Sample size	Combination matrix ($H = 5$)								
	$\mathbf{C}_1$			$\mathbf{C}_2$			$\mathbf{C}_3$		
	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$
20	.16	.10	.04	.21	.15	.06	.13	.08	.01
50	.10	.06	.02	.12	.06	.02	.10	.06	.01
100	.09	.05	.01	.09	.05	.01	.09	.05	.01
200	.09	.05	.01	.10	.05	.01	.10	.05	.01

Proportion of $\overline{\mathbf{K}}_{\text{calc}}$ values exceeding the $F_{M,T-Mp-1}(\alpha)$ quantile with $\Sigma^*_{\mathbf{u}}$ obtained with $N = 1000$ Monte Carlo replications of (2.14)–(2.15).

tistic produces p_α values closer to the nominal ones. However, the improvement may not be enough for some applications. This should be taken into account when making conclusions from a restricted forecasting analysis.

2.4.4 A simulation study for the Mexican data example

An empirical application for the Mexican Economy is presented in the following section to illustrate the methodology. The VAR(3) model employed includes six variables plus a constant and some dummy variables. Since the system turned out to be CI$(1,1)$, the estimation was done with a VEC representation. After adjusting by the order of the process the sample size is equal to $T = 52$.

A numerical algorithm similar to that described above was used for the VAR process whose estimated results appear in Section 2.5.2 with $N = 1000$, $T = 52$, $p = 3$, the $\mathbf{C}$ matrix given in (2.23), $M = 3$, and $H = 8$. Table 2.5 shows the p_α proportions of $\mathbf{K}$ and $\overline{\mathbf{K}}$ for $\Sigma_{\mathbf{u}} = \mathbf{0}$ and $\Sigma_{\mathbf{u}}^*$.

It is clear that $\overline{\mathbf{K}}$ has better performance than $\mathbf{K}$, even though the p_α values for $\overline{\mathbf{K}}$ with $\Sigma_{\mathbf{u}}^*$ are much smaller than α. Apparently the uncertainty distorts the distribution of the JCTs for small samples. Nevertheless, the p_α values for $\overline{\mathbf{K}}$ with $\Sigma_{\mathbf{u}}^*$ are closer to α than those for $\mathbf{K}$. So, the feasible JCT performs better than the JCT for the VEC model of the Mexican economic system.

Table 2.5: p_α values for the estimated Mexican system

Covariance matrix ($H = 8$)	$\Sigma_{\mathbf{u}} = \mathbf{0}$			$\Sigma_{\mathbf{u}}^*$		
Statistic	$p_{.10}$	$p_{.05}$	$p_{.01}$	$p_{.10}$	$p_{.05}$	$p_{.01}$
$\mathbf{K}$	.41	.32	.18	.26	.18	.07
$\overline{\mathbf{K}}$	.10	.04	.01	.04	.01	.00

Proportions of $\mathbf{K}$ and $\overline{\mathbf{K}}$ for $\Sigma_{\mathbf{u}} = \mathbf{0}$ and $\Sigma_{\mathbf{u}}^*$ values exceeding their theoretical quantiles for the estimated VEC of the Mexican economy with restrictions given by (2.23) and forecast horizon of $H = 8$. The Monte Carlo experiment was done for $N = 1000$ replications.

2.5 Empirical illustration

At the end of 2002 the Mexican Government published the economic targets for 2003 (see SHCP, 2002). There, it was foreseen that the rate of growth of GDP would move from 1.7% in 2002 to 3.0% in 2003. The annual inflation rate was targeted to be reduced from 4.9% in 2002 to 3.0% in 2003 and the trade balance deficit was supposed to move from -15,234.6 (-2.4% of GDP in 2002) to -18,035.5 (-2.8% of GDP in 2003).

The data set consists of 55 quarterly observations. When the Mexican Government announced the targets the data available ran up to 2002:III. So, the estimation period covers data from 1989:I to 2002:III. The data are available from the web site at ITAM with the address http://allman.rhon.itam.mx/%7E guerrero/Series_VEC.pdf. A description of the variables employed follows.

Gross domestic product (LGDP). Measured in thousands of Mexican pesos at constant prices of 1993. Source: *Instituto Nacional de Estadística Geografía e Informática, Sistema Nacional de Cuentas Nacionales*. The series is log transformed, $\mathrm{LGDP}_t = \ln(\mathrm{GDP}_t)$.

Mexican inflation rate (PMEX). First difference of log Consumer Price Index (ipcmex) with base 1994=100. Source: *Bank of Mexico*. The Consumer Price Index is monthly and the quarterly series are the values at the end of the quarters. $\mathrm{PMEX}_t = \ln(\mathrm{ipcmex}_t) - \ln\left(\mathrm{ipcmex}_{t-1}\right)$.

Unemployment rate (LUNMP). Source: *Instituto Nacional de Estadística Geografía e Informática, National Urban Employment Survey.* The data are log transformed, $\mathrm{LUNMP}_t = \ln(\mathrm{UNMP}_t)$.

Real demand of money (LMONB). Currency held by the public plus domestic currency and checking accounts in resident banks. This is a monthly series given in nominal terms in thousand of Mexican pesos (basemon). The quarterly series is obtained by averaging the monthly values and is deflated by ipcmex. Source: *Bank of Mexico.* The series is log transformed, $\mathrm{LMONB}_t = \ln(\mathrm{basemon}_t/\mathrm{ipcmex}_t)$.

Trade balance deficit (TRDB). Defined as income minus expenditure of the foreign sector. This is a quarterly series given in millions of dollars (DEF). Source: *Bank of Mexico.* The series is transformed by dividing it by 10,000 to homogenize the data, $\mathrm{TRDB}_t = \mathrm{DEF}_t/10{,}000$.

US inflation rate (PUSA). First difference of the log US Consumer Price Index (ipcusa) with base 1982-84=100. Source: *US Department of Labor, Bureau of Labor Statistics*. The US Consumer Price Index is monthly and the quarterly series are the values at the end of the quarters, $\text{PUSA}_t = \ln(\text{ipcusa}_t) - \ln(\text{ipcusa}_{t-1})$.

2.5.1 Order of integration

The order of integration of the series was decided by Augmented Dickey-Fuller (ADF) tests. The augmented regression model included a constant, centered dummies for PMEX_t, LUNMP_t, TRDB_t and PUSA_t and a deterministic trend for LGDP_t and LMONB_t. The general equation was

$$\Delta y_t = \alpha + \gamma_0 t + \sum_{i=1}^{3} \gamma_i D_{it} + \delta_0 y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + error$$

In order to account for the Mexican crisis of 1995 we included two dummy variables for the first and second quarters of that year. Table 2.6 shows the results of the ADF tests with and without dummy variables. That is, with and without accounting for structural change (SC).

The order of the autoregression p was selected to guarantee no autocorrelation of the residuals. The τ statistic allows one to test $H_0 : \delta_0 = 0$. Critical values do not consider intervention dummy variables to account for the 1995 crisis. Except for domestic inflation (PMEX) in levels, the order of integration of the variables did not depend on the inclusion of the dummy variables, but there still was some

Table 2.6: ADF unit root test results

Variable	$H_0 : I(1)$				$H_0 : I(2)$			
	Without SC		With SC		Without SC		With SC	
	p	τ	p	τ	p	τ	p	τ
LGDP	0	-2.16	0	-3.06	0	$-6.50*$	0	$-10.21*$
PMEX	3	-2.43	0	$-5.70*$	1	$-7.76*$	0	$-9.34*$
LUNMP	1	-1.43	1	-2.42	0	$-5.52*$	0	$-7.12*$
LMONB	1	-1.68	2	-0.25	0	$-4.75*$	2	$-6.61*$
TRDB	0	-2.18	4	-1.66	0	$-6.56*$	1	$-7.43*$
PUSA	3	-2.81	–	–	1	$-9.40*$	–	–

Note: * indicates rejection of H_0 at the 5% significance level.

doubt whether or not inflation is stationary. However, since the test result can be distorted by the inclusion of dummy variables, we assumed that PMEX is $I(1)$.

2.5.2 VEC estimation

The VAR model included the six economic variables previously described, $\mathbf{y}_t = (\text{PMEX}_t, \text{LGDP}_t, \text{LMONB}_t, \text{TRDB}_t, \text{LUNMP}_t, \text{PUSA}_t)'$, a constant, centered dummy variables to account for seasonal effects and two dummy variables to account for the 1995 crisis, i.e. $\mathbf{D}_t = \left(\text{const}, S_{1,t}, S_{2,t}, S_{3,t}, I_{95:I,t}, I'_{95:II,t}\right)$. Although the US inflation rate is basically exogenous to the Mexican economy, it entered the model as an endogenous variable because in the estimation stage it was found that not other variable in the system affects PUSA significantly. The system became CI$(1,1)$ and an integrated VAR(3) model provided a reasonable representation for the system in levels. Therefore the system was estimated with the following VEC model

$$\Delta \mathbf{y}_t = \Lambda \mathbf{D}_t + \alpha\beta' \mathbf{y}_{t-1} + \Pi_1^* \Delta \mathbf{y}_{t-1} + \Pi_2^* \Delta \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t$$

The estimation results are as follows (t-values in parentheses)

$$\widehat{\alpha} = \begin{pmatrix} \underset{(-1.71)}{-0.279} & \underset{(3.48)}{0.642} & \underset{(-1.21)}{-0.467} & \underset{(-1.28)}{-1.607} & \underset{(-1.82)}{-2.659} & \underset{(0.79)}{0.049} \end{pmatrix},$$

$$\widehat{\beta} = \begin{pmatrix} 1 & -0.038 & 0.058 & -0.036 & 0.017 & -0.221 \end{pmatrix},$$

$$\widehat{\Lambda} = \begin{pmatrix} - & - & \underset{(-2.13)}{-0.028} & - & \underset{(9.02)}{0.113} & \underset{(5.30)}{0.102} \\ - & \underset{(-3.63)}{-0.071} & \underset{(-3.91)}{-0.057} & \underset{(-7.27)}{-0.084} & \underset{(-4.23)}{-0.060} & \underset{(-3.84)}{-0.084} \\ \underset{(2.38)}{0.018} & \underset{(-2.74)}{-0.112} & \underset{(-5.30)}{-0.162} & \underset{(-5.95)}{-0.143} & \underset{(-4.84)}{-0.144} & \underset{(-2.96)}{-0.135} \\ - & \underset{(2.59)}{0.348} & \underset{(2.13)}{0.213} & - & \underset{(6.57)}{0.637} & \underset{(2.50)}{0.372} \\ - & - & \underset{(2.02)}{0.235} & - & \underset{(3.12)}{0.351} & - \\ - & \underset{(2.60)}{0.017} & - & \underset{(3.00)}{0.012} & - & - \end{pmatrix},$$

$$\widehat{\Pi}_1^* = \begin{pmatrix} \underset{(-2.16)}{-0.350} & - & - & - & - & - \\ \underset{(-3.20)}{-0.588} & \underset{(-3.56)}{-0.525} & \underset{(3.13)}{0.265} & \underset{(2.03)}{0.046} & - & \underset{(2.32)}{1.193} \\ - & - & - & - & - & - \\ - & - & - & \underset{(-3.57)}{-0.548} & \underset{(2.27)}{0.341} & \underset{(2.50)}{8.749} \\ \underset{(2.61)}{3.795} & - & - & - & - & - \\ - & - & - & - & - & \underset{(-1.97)}{-0.335} \end{pmatrix},$$

$$\widehat{\Pi}_2^* = \begin{pmatrix} \underset{(-2.22)}{-0.225} & - & - & - & - & \underset{(2.09)}{0.804} \\ \underset{(-1.95)}{-0.224} & - & - & - & - & - \\ - & - & - & - & - & - \\ - & - & - & \underset{(-2.42)}{-0.313} & - & \underset{(2.76)}{8.230} \\ - & - & - & - & - & - \\ - & - & - & - & - & - \end{pmatrix},$$

with R^2 equal to 0.89, 0.96, 0.92, 0.78, 0.70, and 0.81 for ΔPMEX, ΔLGDP, ΔLMONB, ΔTRDB, ΔLUNMP and ΔPUSA, respectively. The matrices $\widehat{\Lambda}$, $\widehat{\Pi}_1^*$, and $\widehat{\Pi}_2^*$ show only the numerical values of those elements found significant at the 5% level.

The following matrix shows the contemporaneous residual correlations. By symmetry, only the lower diagonal part is shown.

$$
\begin{array}{c c}
 & \begin{array}{cccccc} \text{PMEX} & \text{LGDP} & \text{LMONB} & \text{TRDB} & \text{LUNMP} & \text{PUSA} \end{array} \\
\begin{array}{c} \text{PMEX} \\ \text{LGDP} \\ \text{LMONB} \\ \text{TRDB} \\ \text{LUNMP} \\ \text{PUSA} \end{array} &
\begin{bmatrix}
1.00 & & & & & \\
0.15 & 1.00 & & & & \\
-0.39 & 0.09 & 1.00 & & & \\
-0.08 & 0.01 & 0.23 & 1.00 & & \\
-0.32 & -0.11 & -0.06 & -0.17 & 1.00 & \\
0.24 & 0.09 & -0.22 & 0.27 & -0.19 & 1.00
\end{bmatrix}
\end{array}
$$

There is negative correlation between Mexican inflation rate and real demand for money, as it should be expected. There is also positive correlation between US inflation rate and trade balance deficit. In Table 2.7 we report the results of Johansen tests (see for instance Johansen, 1988).

At the 5% significance level there are two cointegrating relationships and only one at the 1% level. We decided to use only one cointegration relationship. The resulting equation was

$$\epsilon_t = \text{PMEX}_t - 0.038\text{LGDP}_t + 0.058\text{LMONB}_t - 0.036\text{TRDB}_t + 0.017\text{LUNMP}_t - 0.221\text{PUSA}_t.$$

whose graph is shown in Fig. 2.1. The observed and fitted series of the VEC model as well as their residuals are shown in Fig. 2.2.

Table 2.7: Johansen cointegration analysis

NULL:	Trace Statistic	Crit 95%	Crit 99%	Eigen Statistic	Crit 95%	Crit 99%
$r \leq 0$	124.23**	95.75	104.96	46.58**	40.08	45.87
$r \leq 1$	77.65*	69.82	77.82	38.34*	33.88	39.37
$r \leq 2$	39.31	47.86	54.68	20.12	27.59	32.72
$r \leq 3$	19.19	29.80	35.46	14.63	21.13	25.87
$r \leq 4$	4.56	15.49	19.94	4.41	14.26	18.52
$r \leq 5$	0.15	3.84	6.64	0.15	3.84	6.64

Note: * rejection at 5% and ** indicates rejection at 1% level.

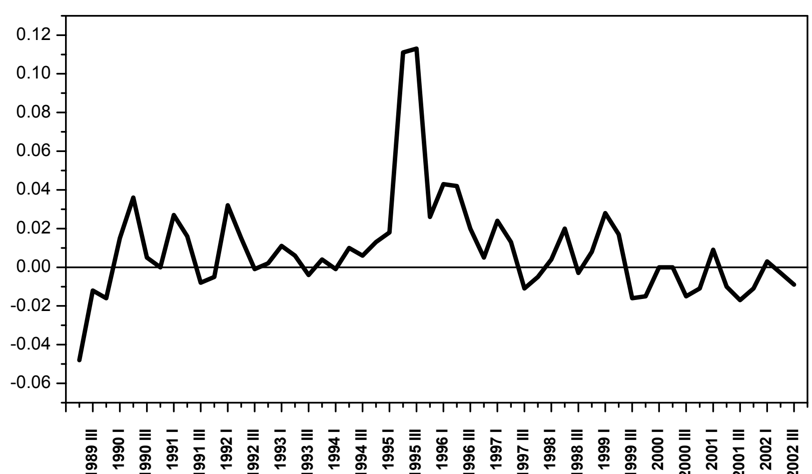


Figure 2.1: Cointegration relationship

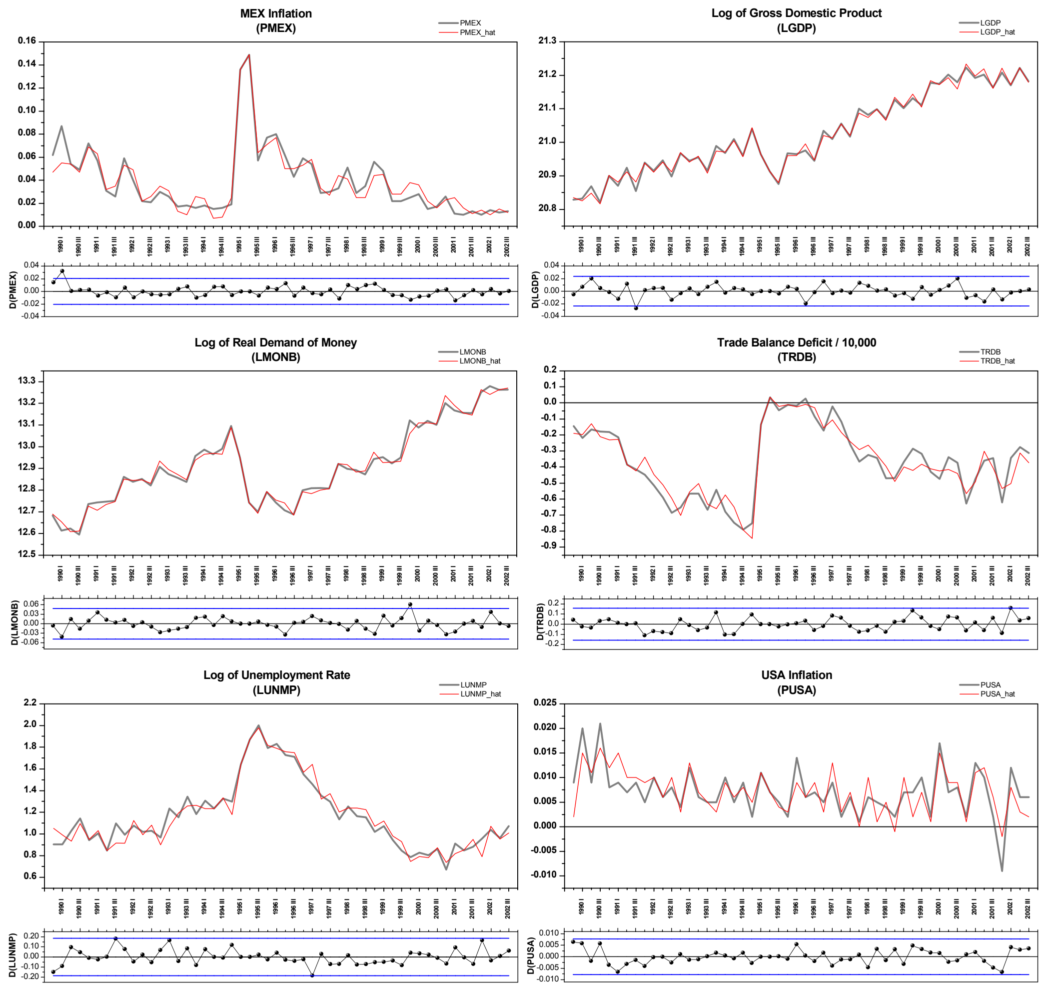


Figure 2.2: Observed and estimated series in levels and corresponding residuals of first differences, with $\pm 2\sigma$ bands plotted as horizontal lines.

2.5.3 Forecasts restricted to fulfill targets

The economic targets described above are written as

$$\mathrm{LGDP}_{2003:IV} - \mathrm{LGDP}_{2002:IV} = \ln(1.03)$$

$$\mathrm{PMEX}_{2003:I} + \mathrm{PMEX}_{2003:II} + \mathrm{PMEX}_{2003:III} + \mathrm{PMEX}_{2003:IV} = \ln(1.03)$$

$$\mathrm{TRDB}_{2003:I} + \mathrm{TRDB}_{2003:II} + \mathrm{TRDB}_{2003:III} + \mathrm{TRDB}_{2003:IV} = -18,035.5/10,000$$

Thus, the linear stochastic restriction (2.7) considers an $H = 8$ period-ahead forecast and gets specified by

$$\mathbf{R} = \begin{pmatrix} \ln(1.03) \\ \ln(1.03) \\ -18,035.5/10,000 \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} -\mathbf{e}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{e}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{2.23}$$

where $\mathbf{0} = (0, ..., 0)$ is the 1×6 zero vector and $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)$ is a 1×6 vector with 1 in the ith position and zeros elsewhere. The restricted forecasts are obtained for both the binding case ($\Sigma_{\mathbf{u}} = \mathbf{0}$) and for the unbinding case with $\Sigma^*_{\mathbf{u}}$.

Table 2.8 reports the feasible JCT for $\Sigma_{\mathbf{u}} = \mathbf{0}$ and $\Sigma^*_{\mathbf{u}}$ as well as their SCTs. The $p-$values were obtained from the $F_{3,42}$ distribution. Neither the binding nor the unbinding restrictions were rejected by the compatibility tests at the usual significance levels. In the unbinding restriction case it can be verified that $3\overline{\mathbf{K}} = \overline{\mathbf{K}}_1 + \overline{\mathbf{K}}_2 + \overline{\mathbf{K}}_3$. Even though none of the single restrictions is incompatible, it is interesting to see that the restriction on inflation produces an SCT statistic $\left(\overline{\mathbf{K}}_2 = 0.19\right)$ that contributes with 53% to the feasible JCT value $\left(3\overline{\mathbf{K}} = 0.36\right)$. This result may be interpreted as saying that the inflation restriction could be the main cause of incompatibility, in case that such incompatibility were significant, which in fact was not.

Table 2.8: Feasible compatibility tests for Mexican economic targets

Covariance matrix	$\Sigma_{\mathbf{u}} = \mathbf{0}$		$\Sigma^*_{\mathbf{u}}$	
Statistic	Value	p-value	Value	p-value
$\overline{\mathbf{K}}_1$	0.12	0.73	0.10	0.76
$\overline{\mathbf{K}}_2$	0.30	0.59	0.19	0.66
$\overline{\mathbf{K}}_3$	0.08	0.77	0.07	0.79
$\overline{\mathbf{K}}$	0.15	0.93	0.12	0.95

Single and joint feasible compatibility tests for the Mexican economic targets in 2003, with $\Sigma_{\mathbf{u}} = \mathbf{0}$ and $\Sigma^*_{\mathbf{u}}$. Targets in 2003: 3.0% rate of growth for GDP, trade balance deficit of $-15,234.6$ (-2.4% of GDP) and 3.0% annual inflation rate.

The estimated matrix employed for getting additivity and its eigenvalues are

$$\Sigma^*_{\mathbf{u}} = \begin{pmatrix} 0.0006 & -0.0002 & 0.0024 \\ -0.0002 & 0.0001 & -0.0013 \\ 0.0024 & -0.0013 & 0.3728 \end{pmatrix} \text{ and } \Lambda = \begin{pmatrix} 0.3728 \\ 0.0006 \\ 0.0000 \end{pmatrix}.$$

Fig. 2.3 shows the unrestricted and binding restricted ($\Sigma_{\mathbf{u}} = \mathbf{0}$) forecast paths with their 90% probability intervals for each of the six economic variables. The corresponding paths for the unbinding restricted $(\Sigma_{\mathbf{u}}^{*})$ forecasts are shown in Fig. 2.4.

Table 2.9 summarizes the 2003 Mexican Government economic targets, the unrestricted forecast, the binding restricted forecast $(\Sigma_{\mathbf{u}} = \mathbf{0})$ and the unbinding restricted forecast $(\Sigma_{\mathbf{u}}^{*})$. Here, we can appreciate that the Government is taking into account the effects of its future monetary policies. For instance, the unrestricted forecast leads to an economic growth and inflation rate of 4.01% and 5.28% respectively. Thus, the monetary policy that controls inflation to stay around 3.00% yields a decrease in economic growth. Notice that the binding restricted forecasts attain the targets exactly. These values are relaxed by the unbinding restricted forecasts if uncertainty is allowed.

Table 2.9: Forecasts of the Mexican system

Variables	Targets	Forecasts		
		Unrestricted	Restricted with $\Sigma_{\mathbf{u}} = \mathbf{0}$	Restricted with $\Sigma_{\mathbf{u}}^{*}$
GDP	3.00%	4.01%	3.00%	3.14%
PMEX	3.00%	5.28%	3.00%	3.53%
UNMP	–	2.63%	2.68%	2.66%
MONB	–	9.47%	9.92%	9.67%
DEF	$-18,035.5$	$-16,701.4$	$-18,035.5$	$-17,730.7$

VEC forecasts results for the Mexican system in 2003.

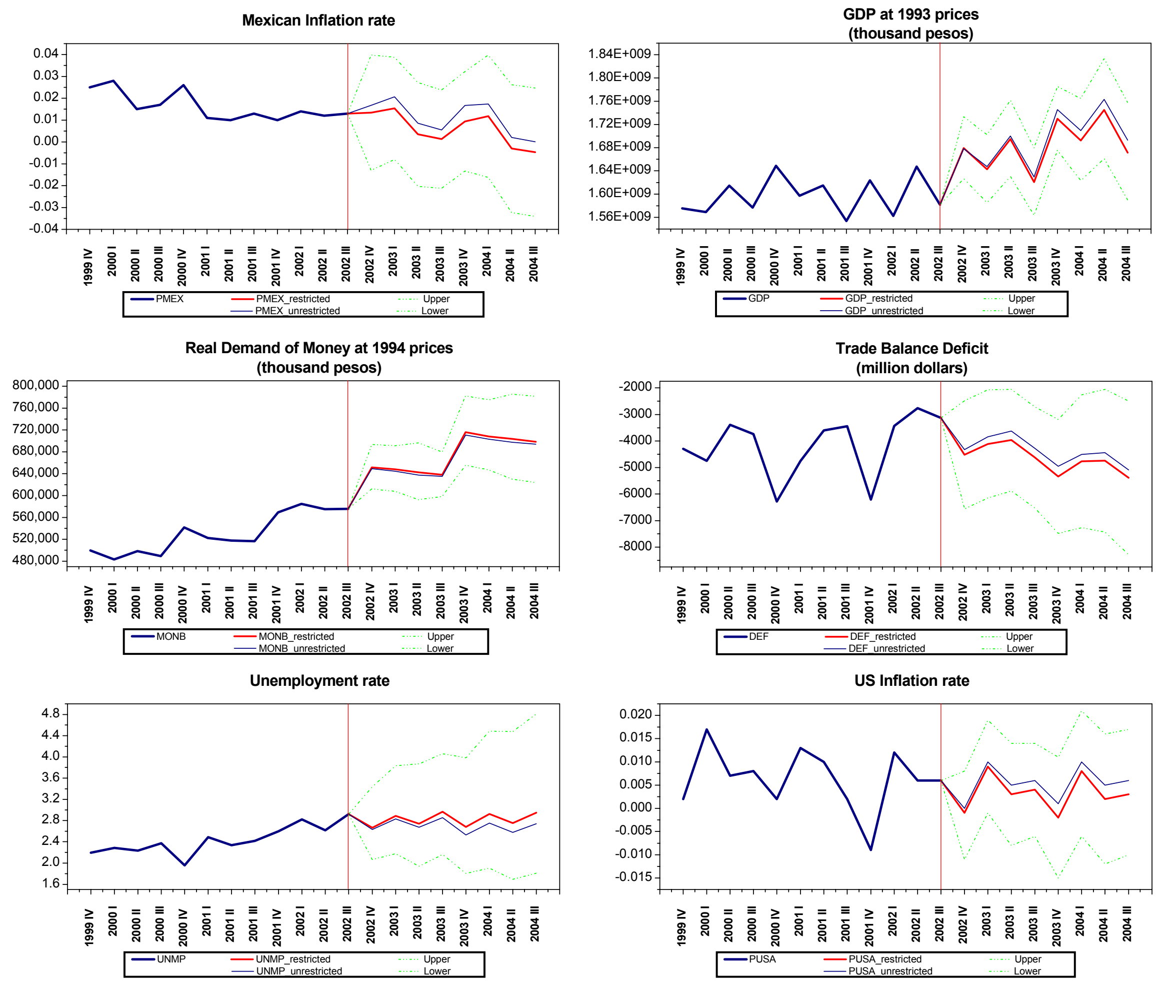


Figure 2.3: Restricted forecasts with 90% probability intervals for $\Sigma_u = \mathbf{0}$ and unrestricted forecasts (origin at 2002:III).

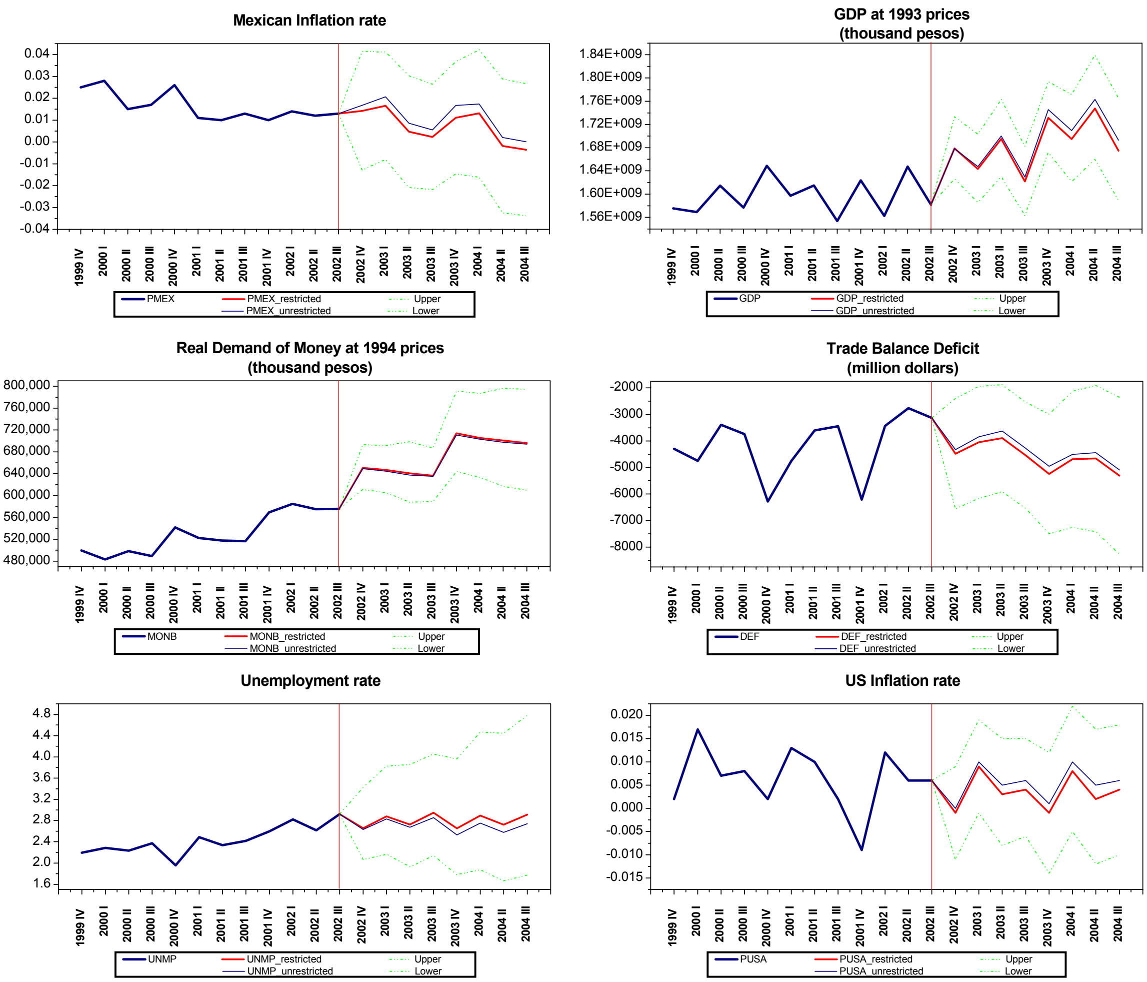


Figure 2.4: Restricted forecasts with 90% probability intervals for Σ_u^* and unrestricted forecasts (origin at 2002:III).

2.5.4 Forecasts restricted by unrealistic targets

Since the Mexican economic targets for 2003 turned out to be compatible (2.5.3), that is, neither the JCT nor the SCTs provided evidence of incompatibility, it is difficult to appreciate the relevance of using $\Sigma_{\mathbf{u}}^{*}$ to identify which restrictions were the most likely causes of rejection of the JCT. We therefore consider an unrealistic case in which the targets are: 10.5% rate of growth for GDP in 2003, trade balance deficit of $-32,151$ (-5.0% of GDP) and 3.0% annual inflation rate. Table 10 reports the feasible JCT and the corresponding SCTs for this situation. Considering a significance level of 10% we should reject the joint compatibility hypothesis for the binding case ($\Sigma_{\mathbf{u}} = \mathbf{0}$). As we can see, the sources of incompatibility are GDP and TRDB, but at this point we should be cautious when interpreting these results because of the lack of additivity of the SCTs with respect to the JCT. On the other hand, additivity can be obtained at the cost of introducing uncertainty in the sense of (2.12) yielding a JCT value lower than that obtained with $\Sigma_{\mathbf{u}} = \mathbf{0}$. Thus, some binding restrictions which are incompatible may become compatible when introducing uncertainty. However, introducing uncertainty by $\Sigma_{\mathbf{u}}^{*}$ does not always make the restrictions compatible. This is the case reported in Table 2.10, where the restriction on GDP becomes compatible, while the restriction on TRDB remains incompatible. The joint compatibility test also rejects compatibility. Of course, these results depend on the significance level employed. Finally, from the additivity of the JCT with $\Sigma_{\mathbf{u}}^{*}$ we can see that the main source of incompatibility is the TRDB restriction, which amounts to 58% of the JCT value.

Table 2.10: Feasible compatibility tests for unrealistic economic targets

Covariance matrix	$\Sigma_{\mathbf{u}} = \mathbf{0}$		$\Sigma^*_{\mathbf{u}}$	
Statistic	Value	p-value	Value	p-value
$\overline{\mathbf{K}}_1$	4.65	0.04	2.69	0.11
$\overline{\mathbf{K}}_2$	0.30	0.59	0.29	0.60
$\overline{\mathbf{K}}_3$	11.27	0.00	4.08	0.05
$\overline{\mathbf{K}}$	4.71	0.01	2.35	0.09

Single and joint feasible compatibility tests for unrealistic economic targets in 2003, with $\Sigma_{\mathbf{u}} = \mathbf{0}$ and $\Sigma^*_{\mathbf{u}}$. Unrealistic economic targets in 2003: 10.5% rate of growth for GDP, trade balance deficit of $-32,151$ (3.0% annual inflation rate.

2.6 Conclusions

This paper presents a complement to the multivariate restricted forecasting methodology to study compatibility between targets and unrestricted forecasts from a VEC model. The compatibility test can be used to detect structural breaks during the forecast horizon and to monitor the attainment of economic targets. The restrictions can be binding or unbinding. In the latter case the uncertainty may come from subject matter knowledge or, as it is proposed here, can be deduced from the very data. We propose to compute the covariance matrix associated with the restrictions, so that it cancels out model dynamics and interactions between restrictions and thus obtain additivity of the SCTs with respect to the JCT, perhaps at the cost of inducing compatibility. This fact may be considered a drawback of our proposal.

Since the compatibility test for multivariate series relies on asymptotic theory, a numerical simulation was performed to check the small sample performance of the JCT. Then, a feasible JCT that takes into account estimated parameters was obtained and its finite sample properties were also studied. In general, for finite sample sizes the feasible JCT proved to perform better than the original JCT.

The proposed methodology was illustrated with a six-variable Mexican economic system with quarterly data. Implementation of the model focussed on the economic targets for GDP, inflation rate and trade balance deficit for 2003. A numerical simulation of this system was carried out to validate the use of the feasible JCT in this situation. It turned out that the economic targets were compatible with the unrestricted forecasts. Then, some unrealistic targets were considered to illustrate an incompatibility situation.

2.7 Appendices

2.7.1 Appendix A. Approximation of MSE($\mathbf{Y}_{F,H}$) for an estimated process

To derive an expression for $\widetilde{\Sigma}$, the derivatives of $\partial\mathbf{Y}_{F,H}/\partial\boldsymbol{\beta}'$ are needed. They can be obtained by noticing that expression (2.4) can be written as $\mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \mathbf{J}\mathfrak{B}^h\mathbf{Z}_T$, where

$$\mathbf{Z}_T = \begin{pmatrix} \mathbf{D}'_{T+H} & \mathbf{D}'_{T+H-1} & \cdots & \mathbf{D}'_{T+1} & \mathbf{y}'_T & \mathbf{y}'_{T-1} & \cdots & \mathbf{y}'_{T-p+2} & \mathbf{y}'_{T-p+1} \end{pmatrix}'$$

is an $(nH + kp) \times 1$ matrix of variables

$$\mathfrak{B} = \begin{pmatrix} \mathbf{0}_n & & & & & & & & \\ \mathbf{I}_n & \mathbf{0}_n & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & \mathbf{I}_n & \mathbf{0}_n & & & & \\ & & & & \Lambda & \Pi_1 & \cdots & \Pi_{p-1} & \Pi_p \\ & & & & & \mathbf{I}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\ & & & & & & \ddots & \vdots & \vdots \\ & & & & & & & \mathbf{I}_k & \mathbf{0}_k \end{pmatrix},$$

is an $(nH + kp) \times (nH + kp)$ matrix of coefficients and

$$\mathbf{J} = \begin{pmatrix} \mathbf{0}_{k\times n} & \mathbf{0}_{k\times n} & \cdots & \cdots & \mathbf{0}_{k\times n} & \mathbf{I}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k \end{pmatrix}$$

is a $k \times (nH + kp)$ matrix.

Since

$$\begin{aligned}\frac{\partial \mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y})}{\partial \boldsymbol{\beta}'} &= (\mathbf{Z}_T' \otimes \mathbf{J}) \left[\sum_{i=0}^{h-1} (\mathfrak{B}')^{h-1-i} \otimes \mathfrak{B}^i\right] \frac{\partial \mathrm{vec}(\mathfrak{B})}{\partial \boldsymbol{\beta}'} \\ &= \left[\sum_{i=0}^{h-1} \mathbf{Z}_T' (\mathfrak{B}')^{h-1-i} \otimes \mathbf{J}\mathfrak{B}^i\right] \frac{\partial \mathrm{vec}(\mathfrak{B})}{\partial \boldsymbol{\beta}'}\end{aligned}$$

After some algebraic calculations, it can be shown that

$$\frac{\partial \mathrm{vec}(\mathfrak{B})}{\partial \boldsymbol{\beta}'} = \left(\widetilde{\mathbf{I}} \otimes \mathbf{J}'\right) \text{ and } \widetilde{\mathbf{I}} = \begin{pmatrix} \mathbf{0}_{(H-1)n \times (n+kp)} \\ \\ \mathbf{I}_{(n+kp)\times(n+kp)} \end{pmatrix}$$

are $k \times k(n+p)$ and $(Hn + kp) \times (n + kp)$ matrices. Now, we can write

$$\frac{\partial \mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y})}{\partial \boldsymbol{\beta}'} = \sum_{i=0}^{h-1} \mathbf{Z}_T' (\mathfrak{B}')^{h-1-i} \widetilde{\mathbf{I}} \otimes \mathbf{J}\mathfrak{B}^i \mathbf{J}'.$$

From this we have that

$$\frac{\partial \mathbf{Y}_{F,H}}{\partial \boldsymbol{\beta}'} = \begin{pmatrix} \mathbf{Z}_T' \widetilde{\mathbf{I}} \otimes \mathbf{J}\mathbf{J}' \\ \\ \sum_{i=0}^{1} \mathbf{Z}_T' (\mathfrak{B}')^{1-i} \widetilde{\mathbf{I}} \otimes \mathbf{J}\mathfrak{B}^i \mathbf{J}' \\ \\ \vdots \\ \\ \sum_{i=0}^{H-1} \mathbf{Z}_T' (\mathfrak{B}')^{H-1-i} \widetilde{\mathbf{I}} \otimes \mathbf{J}\mathfrak{B}^i \mathbf{J}' \end{pmatrix}.$$

Let $Y_t = \left(1, \mathbf{y}_t', \ldots, \mathbf{y}_{t-p+1}'\right)'$ be a $(kp+1) \times 1$ vector and $Z = (Y_0, \ldots, Y_{T-1})$ a $(kp+1) \times T$ matrix. Thus the asymptotic covariance matrix of $\widehat{\boldsymbol{\beta}}$ can be written as $\Sigma_{\widehat{\boldsymbol{\beta}}} = \Gamma^{-1} \otimes \Sigma_{\boldsymbol{\varepsilon}}$ where $\Gamma \equiv \mathrm{plim} ZZ'/T$ and $\Sigma_{\boldsymbol{\varepsilon}}$ is the error covariance matrix. Thus expression (2.21) becomes

$$\widetilde{\Sigma} = \mathrm{E}\left[\frac{\partial \mathbf{Y}_{F,H}}{\partial \boldsymbol{\beta}'} \left(\Gamma^{-1} \otimes \Sigma_{\boldsymbol{\varepsilon}}\right) \frac{\partial \mathbf{Y}_{F,H}'}{\partial \boldsymbol{\beta}}\right] = \left\{\widetilde{\Sigma}_{m,n}\right\} \text{ for } m, n = 1, \ldots, H,$$

where

$$\widetilde{\Sigma}_{m,n} = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \mathrm{tr}\left[(\mathfrak{B}')^{m-1-i}\widetilde{\mathbf{I}}\Gamma^{-1}\widetilde{\mathbf{I}}'\mathfrak{B}^{n-1-j}\mathrm{E}\left(\mathbf{Z}_T\mathbf{Z}_T'\right)\right]\Psi_i\Sigma_{\boldsymbol{\varepsilon}}\Psi_j' \text{ for } m,n = 1,\ldots,H,$$

are $k \times k$ matrices and $\Psi_i = \mathbf{J}\mathfrak{B}^i\mathbf{J}'$.

2.7.2 Appendix B. Restricted forecasts for estimated process

In this appendix we obtain the restricted forecast expression for the VAR model with estimated coefficients $\widehat{\Theta}$. From the linear projection theory we know that

$$\widehat{\mathbf{Y}}_{F,H}^R = \widehat{\mathbf{Y}}_{F,H} + \widehat{\mathrm{E}}\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}\middle| \mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right),$$

where $\widehat{\mathbf{Y}}_{F,H}^R \equiv \widehat{\mathrm{E}}\left(\mathbf{Y}_F|\mathbf{Y},\mathbf{R}\right)$ is the restricted forecast for an estimated process and $\mathbf{Y}_F - \widehat{\mathrm{E}}(\mathbf{Y}_F|\mathbf{Y})$ is orthogonal to $\mathbf{Y}$ conditional on $\mathbf{R} - \widehat{\mathrm{E}}(\mathbf{R}|\mathbf{Y})$. Now

$$\widehat{\mathrm{E}}\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}\middle| \mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right) = \widehat{\mathbf{A}}\left(\mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right)$$

for some $\widehat{\mathbf{A}}$ which is an $H \times M$ matrix. By substituting (2.18) and (2.22) into the following orthogonality condition we get

$$\widehat{\mathrm{E}}\left\{\left[\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}\right) - \widehat{\mathrm{E}}\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}\middle| \mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right)\right]\left(\mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right)\right\} = \mathbf{0}$$

so that

$$\left(\Sigma_{\mathbf{Y}_{F,H}} + T^{-1}\widetilde{\Sigma}\right)\mathbf{C}' - \widehat{\mathbf{A}}\mathbf{C}\left(\Sigma_{\mathbf{Y}_{F,H}} + T^{-1}\widetilde{\Sigma}\right)\mathbf{C}' - \widehat{\mathbf{A}}\Sigma_{\mathbf{u}} = \mathbf{0},$$

and

$$\widehat{\mathbf{A}} = \Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}'\widehat{\boldsymbol{\Omega}}^{-1} \text{ where } \widehat{\boldsymbol{\Omega}} = \mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}' + \Sigma_{\mathbf{u}}.$$

Therefore, the optimal forecast of $\mathbf{Y}_F$ conditional on $\mathbf{Y}$ and $\mathbf{R}$, for an estimated VAR process is

$$\widehat{\mathbf{Y}}_{F,H}^R = \widehat{\mathbf{Y}}_{F,H} + \widehat{\mathbf{A}}\left(\mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right).$$

Its forecast error vector is

$$\begin{aligned}
\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}^R &= \mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H} - \widehat{\mathbf{A}}\left(\mathbf{R} - \mathbf{C}\widehat{\mathbf{Y}}_{F,H}\right) \\
&= \left(I - \widehat{\mathbf{A}}\mathbf{C}\right)\left\{\boldsymbol{\Psi}\boldsymbol{\epsilon}_F + \left[\sqrt{T}\left(\mathbf{Y}_{F,H} - \widehat{\mathbf{Y}}_{F,H}\right)\right]/\sqrt{T}\right\} - \widehat{\mathbf{A}}\mathbf{u},
\end{aligned}$$

hence

$$\begin{aligned}
\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,H}^R\right) &\equiv \mathrm{E}\left[\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}^R\right)\left(\mathbf{Y}_F - \widehat{\mathbf{Y}}_{F,H}^R\right)'\middle|\mathbf{Y}, \mathbf{R}\right] \\
&= \left(I - \widehat{\mathbf{A}}\mathbf{C}\right)\left(\Sigma_{\mathbf{Y}_{F,H}} + T^{-1}\widetilde{\Sigma}\right)\left(I - \mathbf{C}'\widehat{\mathbf{A}}'\right) + \widehat{\mathbf{A}}\Sigma_{\mathbf{u}}\widehat{\mathbf{A}}' \\
&= \left(I - \widehat{\mathbf{A}}\mathbf{C}\right)\Sigma_{\widehat{\mathbf{Y}}_{F,H}}.
\end{aligned}$$

Now, since

$$\mathbf{A}\mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}} = \Sigma_{\widehat{\mathbf{Y}}_{F,H}}\mathbf{C}'\widehat{\boldsymbol{\Omega}}^{-1}\mathbf{C}\Sigma_{\widehat{\mathbf{Y}}_{F,H}}$$

is a symmetric semidefinite positive matrix and $\Sigma_{\widehat{\mathbf{Y}}_{F,H}}$ is the $\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,H}^R\right)$ we have that $\widehat{\mathbf{Y}}_{F,H}^R$ is at least as precise as $\widehat{\mathbf{Y}}_{F,H}$.

BIBLIOGRAPHY

[1] Box, G.E. P. & Tiao, G.C (1976). Comparison of forecasts and actuality, *Applied Statistics,* 25, 195-200.

[2] Clements, M. P. & Hendry, D.F (2004). *A Companion to Economic Forecasting*, Blackwell Publishing.

[3] Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions, *Econometric Reviews,* 3, 1-100.

[4] Dufour, J. M. (1985). Unbiasedness of predictions from estimated vector autoregressions, *Econometric Theory*, 1, 387-402.

[5] Engle, R. F. & Granger, C. W. J. (1987). Co-integration and error correction representation, estimation and testing, *Econometrica,* 55, 251-276.

[6] Greene, M.N., Howrey, E.P. & Hymans, S.W. (1986). The use of outside information in econometric forecasting, *in Model Reliability, eds. E. Huh and D.A. Belsley, Cambridge, MA: MIT Press,* 90-116.

[7] Guerrero, V. M. (1989). Optimal conditional ARIMA forecasts, *Journal of Forecasting,* 8, 215-229.

[8] Guerrero, V.M. & Peña D. (2003). Combining multiple time series predictors: a useful inferential procedure, *Journal of Statistical Planning and Inference,* 116, 249-276.

[9] Guerrero V.M., Pena B., Senra E., & Alegría, A. (2005). Restricted forecasting with a VEC model: Validating the feasibility of economic targets, *Available in Eva Senra's website http://www2.uah.es/ .*

[10] Johansen, S. (1988). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models,* Oxford University Press.

[11] Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.

[12] Magnus, J. R. & Neudecker, H. (2002). *Matrix Differential Calculus*, Wiley.

[13] Nieto, F.H. & Guerrero, V.M. (1995). Kalman filter for singular and conditional state-space models when the system state and the observational error are correlated, *Statistics & Probability Letters,* 22, 303-310.

[14] Pandher, G. S. (2002). Forecasting multivariate time series with linear restrictions using constrained structural state-space models, *Journal of Forecasting,* 21, 281-300.

[15] Pankratz, A. (1989). Time series forecasts and extra-model information, *Journal of Forecasting,* 8, 75-83.

[16] Park, J.Y. & Phillips, P.C.B. (1988). Statistical inference in regression with integrated process: Part 1, *Econometric Theory,* 4, 468 - 497.

[17] Park, J.Y. & Phillips, P.C.B. (1989). Statistical inference in regression with integrated process: Part 2, *Econometric Theory,* 5, 95 - 131.

[18] Trabelsi, A. & Hillmer, S. C. (1989). A benchmarking approach to forecast combination, *Journal of Business and Economic Statistics,* 7, 353 - 362.

[19] SHCP (2002). *Criterios Generales de Política Económica 2003*, México: Secretaría de Hacienda y Crédito Público.

[20] van der Knoop, H.S. (1987). Conditional forecasting with a multivariate time series model. *Economics Letters,* 22, 233 - 236.

CHAPTER 3

RESTRICTED VAR FORECASTS THAT TAKE INTO ACCOUNT AN EXPECTED STRUCTURAL CHANGE

3.1 Introduction

Structural changes are commonly encountered when analyzing time series data. The presence of those extraordinary events in the forecast horizon could easily mislead a time series model and its forecasts, thus leading to erroneous conclusions. For instance, a statistical forecast of the Mexican economy should take into account the structural reform agenda. This happens because nowadays, in Mexico, economic as well as political efforts concentrate on achieving the necessary consensus to advance in the fiscal, energetic and pensions reforms. These reforms are considered necessary to consolidate the macroeconomic stability. If any of these economic reforms were approved, the forecasts based solely on the historical record could mislead the economic expectations. In fact, in a very influential article, Lucas (1976) established that predictions based on historical data would be invalid if some policy change affects the economy, since the economic agents are forward rather than backward-looking and adapt their expectations and behavior to the new policy stance. In this work we consider that an economic reform is likely to occur in the near future and it will cause a structural change in the system. This paper considers the case in which a system of variables are to be forecasted with the aid of a Vector Auto-Regressive (VAR) model. The probability of achieving a given target with the conventional VAR forecast is negligible, unless additional information is taken into account as a restriction on the forecast. We assume here that the expected economic reform will cause a structural change that affects either the deterministic or the stochastic part of the VAR model during the forecast horizon. By doing this we get around Lucas' critique, since the model will not stay the same after the policy change and the forecast will reflect that change. We shall also assume that all the information available on the future effects of the structural change is provided by some economic targets announced by the government. Those targets will be expressed as linear restrictions on the forecasts. Derivation of the restricted forecasts is carried out by Lagrangian optimization. These results generalize those of the univariate case obtained by Guerrero (1991). Furthermore, the univariate time series types of change presented by Tsay (1988) are considered here in a multivariate setting, five of them correspond to deterministic changes and two more serve to capture stochastic changes. Tsay *et al.* (2000) generalized the four most commonly used types of deterministic disturbance effects that appear in univariate time series analysis to the multivariate case.

The problem of incorporating external model information in the univariate time series forecast setting has already been treated in the literature. In fact, Guerrero (1989) and Trabelsi and Hillmer (1989) obtained the optimum restricted forecast, in Mean Squared Error (MSE) sense. The restricted forecast for multiple time series includes diverse works such as those of van der Knopp (1987), Pankratz (1989) and Guerrero *et al.* (2005). In these, the combination of historical information with additional information in the form of linear restrictions, is provided by formulas that consider Vector-Autorregressive and Moving Average (VARMA) models. Within the state-space framework, Pandher (2002) attacked the problem of modelling and forecasting multivariate time series with linear restrictions that apply during the sample period. Such an approach differs from ours in that we consider restrictions that apply in the forecast horizon.

This article is organized as follows. Section 3.2 presents the statistical methodology to get the restricted forecast with VAR models. In section 3.3 we show some typical disturbance functions that are used to model structural changes. In section 3.4 we derive the restricted forecasting formulas of a VAR process affected in both its deterministic and stochastic components. In section 3.5 we illustrate the methodology with an empirical application that uses a VAR model for the Mexican economy. This application assumes that the economic targets for 2004, announced by the government at the end of 2003, will be reached with certainty and that an economic reform that initially will impact GDP and prices, will take place at the beginning of year 2005.

3.2 Methodology

Let $\mathbf{y}_t = (y_{1t} \cdots y_{kt})'$ be a $k \times 1$ vector of variables at time t and let us assume that $\{\mathbf{y}_t\}$ follows a finite pth-order Gaussian VAR model

$$\Pi(B)\mathbf{y}_t = \Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t \tag{3.1}$$

where $\Pi(B) = I - \Pi_1 B - \cdots - \Pi_p B^p$ is a $k \times k$ matrix polynomial of finite degree p and B is the backshift operator such that $B\mathbf{y}_t = \mathbf{y}_{t-1}$. $\boldsymbol{\delta}_t = (\delta_{1t} \cdots \delta_{nt})'$ is an $n \times 1$ vector that includes deterministic variables to account for seasonality as well as intervention effects, and exogenous variables with respect to $\mathbf{y}_t$. $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t} \cdots \varepsilon_{kt})'$ is a $k \times 1$ independent and identically distributed $N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ random error vector with $\boldsymbol{\Sigma}_\varepsilon$ a positive-definite covariance matrix whose i, jth element is $\sigma_{ij} = \mathrm{cov}(\varepsilon_{it}, \varepsilon_{jt})$, for $i, j = 1, 2, \ldots, k$ and $t = 1, 2, \ldots, T$. Thus the $\boldsymbol{\varepsilon}$'s are serially uncorrelated but may be contemporaneously correlated. The effects of $\boldsymbol{\delta}_t$ on $\mathbf{y}_t$ are captured by the $k \times n$ parameter matrix Λ. We assume all the zeros of $\det \Pi(z)$ are on or outside the unit circle.

The VAR model (3.1) can also be written in its moving-average representation form

$$\mathbf{y}_t = \Psi(B)\left(\Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right)$$

where $\Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i$ whose coefficients matrices are given by $\Psi_0 = I$ and $\Psi_i = \sum_{k=1}^{i} \Psi_{i-k}\Pi_k$ ($\Pi_k = 0$ for $k > p$) for $i \geq 1$. Hence, we have $\Psi(B)\Pi(B) = \Pi(B)\Psi(B) = I$.

Subtracting $\mathbf{y}_{t-1}$ from both sides of (3.1) and rearranging terms we have

$$\Pi^*(B)\,\Delta\mathbf{y}_t = \Lambda\boldsymbol{\delta}_t - \Pi(1)\,B\mathbf{y}_t + \boldsymbol{\varepsilon}_t, \tag{3.2}$$

where Δ is the first difference operator and $\Pi^*(B) = I + \sum_{i=1}^{p-1}\Pi_i^* B^i$ is a matrix polynomial of order $p-1$ with $\Pi_i^* = \sum_{j=i+1}^{p}\Pi_j$. Thus $\Pi(B) = \Pi(1)\,B + \Pi^*(B)\,\Delta$.

If $\{\mathbf{y}_t\}$ is unit-root nonstationary ($\det\Pi(1) = 0$) and $\Pi(1) \neq 0$, then there is cointegration in $\{\mathbf{y}_t\}$ and we have that $\Pi(1) = -\boldsymbol{\gamma}\boldsymbol{\beta}$ where $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are $k \times r$ and $r \times k$ matrices with r the rank of $\Pi(1)$, implying that there are $d = k - r$ unit roots in the system. Moreover, we assume the unit roots have multiplicity one, meaning that the elements of $\{\mathbf{y}_t\}$ are at most integrated of order 1, which is written as $I(1)$. When $r > 0$ the variables are cointegrated, in the sense that there exists a linear combination $\boldsymbol{\beta}\mathbf{y}_t$, with $\boldsymbol{\beta} = (\beta_1 \cdots \beta_k) \neq \mathbf{0}$, which is $I(0)$.

Equation (3.2) then becomes

$$\Pi^*(B)\,\Delta\mathbf{y}_t = \Lambda\boldsymbol{\delta}_t + \boldsymbol{\gamma} z_{t-1} + \boldsymbol{\varepsilon}_t, \tag{3.3}$$

where $z_{t-1} = \boldsymbol{\beta}\mathbf{y}_{t-1}$ is stationary. The rows of $\boldsymbol{\beta}$ are then referred as cointegration vectors of the system and equation (3.3) is known as the Vector Error Correction (VEC) representation of (3.1). The term $\Pi^*(B)\,\Delta\mathbf{y}_t$ in this model captures the short-run relationships among the variables, while $\Pi(1)\,B\mathbf{y}_t = -\boldsymbol{\gamma} z_{t-1}$ captures the long-run relationships.

To get the forecast and its MSE of (3.1) let us start by defining $\mathbf{Y} = (\mathbf{y}'_{-p+1} \cdots \mathbf{y}'_T)'$, a $k\,(T+p) \times 1$ vector containing all the past information of the multiple time series and let $\mathbf{Y}_F = (\mathbf{y}'_{T+1} \cdots \mathbf{y}'_{T+H})'$ denote a $kH \times 1$ vector that contains the $H \geq 1$ values to be forecasted for each series. The optimal (in MSE sense) linear forecast of $\mathbf{y}_{T+h}$, for $h = 1, \ldots, H$, is its conditional expectation

$$\mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \Lambda \boldsymbol{\delta}_{T+h} + \Pi_1 \mathrm{E}(\mathbf{y}_{T+h-1}|\mathbf{Y}) + \cdots + \Pi_p \mathrm{E}(\mathbf{y}_{T+h-p}|\mathbf{Y})$$

where $\mathrm{E}(\mathbf{y}_{T+h-i}|\mathbf{Y}) = \mathbf{y}_{T+h-i}$ for $i \geq h$.

Such a forecast produces the forecast error vector

$$\mathbf{y}_{T+h} - \mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \sum_{j=0}^{h-1} \Psi_j \boldsymbol{\varepsilon}_{T+h-j}, \text{ for } h = 1, \ldots, H. \tag{3.4}$$

Stacking all the forecast errors (3.4) we have

$$\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \boldsymbol{\Psi} \boldsymbol{\epsilon}_F$$

where $\boldsymbol{\epsilon}_F = \left(\boldsymbol{\varepsilon}'_{T+1} \cdots \boldsymbol{\varepsilon}'_{T+H}\right)' \sim N\left(\mathbf{0}, I \otimes \boldsymbol{\Sigma}_\varepsilon\right)$ is a $kH \times 1$ random vector, with $\otimes$ the Kronecker product and the $kH \times kH$ matrix $\boldsymbol{\Psi}$ is lower triangular with Ψ_0 in its main diagonal, Ψ_1 in its first subdiagonal, Ψ_2 in the second subdiagonal and so on. The MSE of $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})$ is given by

$$\begin{aligned} \mathrm{MSE}\left[\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})\right] &\equiv \mathrm{E}[(\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}))(\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}))'|\mathbf{Y}], \\ &= \boldsymbol{\Psi}\left(I \otimes \boldsymbol{\Sigma}_\varepsilon\right) \boldsymbol{\Psi}'. \end{aligned}$$

3.3 VAR with structural change

When structural changes occur, the original series $\{\mathbf{y}_t\}$ is disturbed and becomes unobservable. In such a case Lucas' critique (see Lucas, 1976) applies and getting forecasts without taking into account the future effects of those changes is a worthless task. Therefore, we should consider some specific ways in which the model may be modified due to the structural change in the economic system. Following Tsay (1988), let us assume that the observed series $\{\widetilde{\mathbf{y}}_t\}$ follows the model

$$\widetilde{\mathbf{y}}_t = \mathbf{y}_t + f(t) \tag{3.5}$$

where $f(t)$ is the vector disturbance function representing the exogenous effects on $\{\mathbf{y}_t\}$. This function may be deterministic or stochastic depending on the type of disturbances. As a generalization of the univariate case we present seven typical disturbances that may affect a VAR model, five of them correspond to deterministic changes and two more represent stochastic changes.

3.3.1 Deterministic changes

Suppose that a deterministic change (D) is expected to occur at time $t = T + h$ with $1 \leq h \leq H$ (an expected change in the forecast horizon). Using the terminology and notation of Tsay *et al.* (2000) we known that this change can be modeled by the following $k \times 1$ vector disturbance function

$$f(t) = \alpha(B)\boldsymbol{\omega}\xi_t^{(T+h)}$$

where $\alpha(B)$ is a matrix polynomial in B to be defined below. $\boldsymbol{\omega} = (\omega_1 \ldots \omega_K)'$ is the initial $k \times 1$ impact vector of an outlier in the series $\{\mathbf{y}_t\}$ and $\xi_t^{(T+h)}$ is a $k \times 1$

indicator time index $T+h$, that is $\xi_{T+h}^{(T+h)} = 1$ and $\xi_t^{(T+h)} = 0$ if $t \neq T+h$. The function $f(t)$ belongs to the class of intervention models of Box and Tiao (1975). It has proved to have a general form which can be used to describe many dynamic disturbances of a time series. Multiple disturbances can be treated in the same manner by considering a vector disturbance function for each period.

To analyze the effect of cointegration on the outlier dynamics it is convenient to compute the following filtered series $\{a_t\}$

$$a_t = \widetilde{\mathbf{y}}_t - \Lambda \boldsymbol{\delta}_t - \sum_{i=1}^{p} \Pi_i \widetilde{\mathbf{y}}_{t-i} \tag{3.6}$$

where $\widetilde{\mathbf{y}}_t = \mathbf{y}_t$ and $a_t = \boldsymbol{\varepsilon}_t$ for $t < T+h$. In the presence of outliers $a_t \neq \boldsymbol{\varepsilon}_t$ for some t points, otherwise $a_t = \boldsymbol{\varepsilon}_t$ for all t. From (3.1) and writing this last equation as $\Pi(B)\widetilde{\mathbf{y}}_t = \Lambda \boldsymbol{\delta}_t + a_t$ we have

$$a_t = \boldsymbol{\varepsilon}_t + \Pi(B)\alpha(B)\boldsymbol{\omega}\xi_t^{(T+h)}. \tag{3.7}$$

Thus the effect of an outlier on the filter depends on the interaction between $\Pi(B)\alpha(B)$ and $\boldsymbol{\omega}$.

(a) An **innovational outlier** is produced when $\alpha(B) = \Psi(B)$. It represents a change in the innovational series $\{\boldsymbol{\varepsilon}_t\}$ and has a dynamic effect on $\{\mathbf{y}_t\}$ by propagating through the Ψ_i-weights

$$\widetilde{\mathbf{y}}_t = \Psi(B)\left(\boldsymbol{\omega}\xi_t^{(T+h)} + \Lambda \boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right).$$

This innovational change will affect $\mathbf{y}_t$ for $t \geq T+h$. Equation (3.7) becomes $a_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\omega}\xi_t^{(T+h)}$. So, the innovational outlier affects a_t only at time $t = T+h$.

(b) The **additive outlier** will be obtained when $\alpha(B) = I$, so the observed series will be given by

$$\widetilde{\mathbf{y}}_t = \boldsymbol{\omega}\xi_t^{(T+h)} + \Psi(B)\left(\Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right).$$

Notice that the disturbance only affects $\mathbf{y}_{T+h}$.

The defined filtered series are $a_t = \boldsymbol{\varepsilon}_t + \Pi(B)\boldsymbol{\omega}\xi_t^{(T+h)} = \boldsymbol{\varepsilon}_t + \left[\Pi\left(1\right)B + \Pi^*\left(B\right)\Delta\right]\boldsymbol{\omega}\xi_t^{(T+h)}$. So, if $\{\mathbf{y}_t\}$ is a unit-root nonstationary process and $\Pi(1) \neq 0$, then there is cointegration on $\mathbf{y}_t$ and $\Pi(1) = -\boldsymbol{\gamma}\boldsymbol{\beta}$. Let $\boldsymbol{\beta}_\perp$ be a $k \times (k-r)$ full rank orthogonal matrix of $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}\boldsymbol{\beta}_\perp = 0$. The effect of $\boldsymbol{\omega}$ on a_t is as follows.

(i) If $\boldsymbol{\omega}$ is a linear combination of the columns of $\boldsymbol{\beta}_\perp$, then $\Pi(1)\boldsymbol{\omega} = 0$. Thus, the outlier affects the filtered series a_t for $t \geq T+h$ only through the short-run part of the model.

(ii) If $\boldsymbol{\omega}$ is not a linear combination of the columns of $\boldsymbol{\beta}_\perp$, then $\Pi(1)\boldsymbol{\omega} \neq 0$. Therefore, the outlier affects a_t for $t \geq T+h$ both through the short-run and the long-run parts of the model.

(c) A **temporary change** is produced when $\alpha(B) = (1-\rho B)^{-1} I$ with $0 < \rho < 1$. That is

$$\widetilde{\mathbf{y}}_t = (1-\rho B)^{-1}\boldsymbol{\omega}\xi_t^{(T+h)} + \Psi(B)\left(\Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right),$$

this model describes a disturbance $\boldsymbol{\omega}$ that occurs at $t = T+h$ and decays exponentially to the zero vector with rate ρ. This transient change will affect $\{\mathbf{y}_t\}$ for $t \geq T+h$.

Equation (3.7) becomes $a_t = \boldsymbol{\varepsilon}_t + \Pi^*(B)\boldsymbol{\omega}\xi_t^{(T+h)}$, where the coefficient matrices Π_i^* of $\Pi^*(B)$ are $\Pi_i^* = \sum_{j=1}^{i} \rho^{i-j}\Pi_j - \rho^i I$ for $i \leq p$ and $\Pi_i^* = \rho^{i-p}\Pi_p^*$ for $i > p$. Thus $\boldsymbol{\omega}$ affects a_t for $t \geq T+h$ and the effect diminishes with time.

(d) A **level shift** will occur when $\alpha(B) = \Delta^{-1} I$, so that

$$\widetilde{\mathbf{y}}_t = \Delta^{-1}\boldsymbol{\omega}\xi_t^{(T+h)} + \Psi(B)\left(\Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right).$$

This expression says that a level shift $\boldsymbol{\omega}$ occurs in the system at time $t = T+h$ and the change is permanent. Thus, this disturbance will affect $\{\mathbf{y}_t\}$ for $t \geq T+h$.

The effects of cointegration for the level shift case is analyzed in Tsay *et al.* (2000) with the aid of the filtered series. In this case we have $a_t = \boldsymbol{\varepsilon}_t + \Pi^*(B)\boldsymbol{\omega}\xi_t^{(T+h)}$, where the coefficient matrices Π_i^* of $\Pi^*(B)$ are $\Pi_i^* = \sum_{j=1}^{i} \Pi_j - I$ for $i = 1, ..., p$ and $\Pi_i^* = \Pi_p^*$ for $i > p$. In particular we have $\Pi_p^* = -\Pi(1)$. The effect of $\boldsymbol{\omega}$ on a_t is as follows.

(i) If $\boldsymbol{\omega}$ is a linear combination of the columns of $\beta_\perp$, then $\Pi_i^*\boldsymbol{\omega} = -\Pi(1)\boldsymbol{\omega} = 0$ for all $i \geq p$. So the level shift only affects $\{a_t\}$ for $t = 1, ..., p-1$.

(ii) If $\boldsymbol{\omega}$ is not a linear combination of the columns of $\beta_\perp$, then $\Pi_i^*\boldsymbol{\omega} \neq 0$ for all $i \geq p$. Thus, the level shift affects $\{a_t\}$ for $t \geq T+h$.

(e) A **gradual change** is produced when $\alpha(B) = \left[\Delta\left(1-\rho B\right)\right]^{-1} I$ with $0 < \rho < 1$, that is

$$\widetilde{\mathbf{y}}_t = \left[\Delta\left(1-\rho B\right)\right]^{-1}\boldsymbol{\omega}\xi_t^{(T+h)} + \Psi(B)\left(\Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right).$$

This model describes a disturbance $\boldsymbol{\omega}$ that occurs at $t = T + h$ and grows to $(1-\rho)^{-1}\boldsymbol{\omega}$. The gradual change will affect $\{\mathbf{y}_t\}$ for $t \geq T + h$.

The first four of these changes appear in Tsay *et al.* (2000) who generalized the most commonly used deterministic disturbances in univariate time series analysis to the k-dimensional vector autoregressive integrated moving-average (ARIMA) case.

3.3.2 Stochastic changes

Suppose that a stochastic change (V) is expected to occur at time $t = T + h$ and assume that

$$f(t) = \alpha(B)\boldsymbol{\zeta}_t S_t^{(T+h)}$$

where $\{\boldsymbol{\zeta}_t\}$ is a sequence of $k \times 1$ random vectors i.i.d. $N(\mathbf{0}, \boldsymbol{\Sigma}_{\zeta})$ uncorrelated with $\boldsymbol{\varepsilon}_t$, where $\boldsymbol{\Sigma}_{\zeta}$ is the contaminating covariance matrix. $S_t^{(T+h)}$ is the step variable for the time index $T + h$, that is $S_t^{(T+h)} = 1$ for $t \geq T + h$ and zero otherwise. This is a very simple but practical approach to contaminate the original process.

(f) A **variance innovational change** is produced when $\alpha(B) = \Psi(B)$, in which case

$$\widetilde{\mathbf{y}}_t = \Psi(B)\left(\boldsymbol{\zeta}_t S_t^{(T+h)} + \Lambda\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right),$$

this change will affect the variance of $\{\mathbf{y}_t\}$ for $t \geq T+h$. Its univariate counterpart can be found in Tsay (1988) where it is called variance change.

Equation (3.7) becomes $a_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\zeta}_t S_t^{(T+h)}$. So, the variance innovational change makes $a_t \neq \boldsymbol{\varepsilon}_t$ for $t \geq T+h$. What is more, $a_t \sim N(0, \boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_{\boldsymbol{\zeta}})$ for $t \geq T+h$.

(g) The **variance additive change** is produced when $\alpha(B) = I$, in such a way that

$$\widetilde{\mathbf{y}}_t = \boldsymbol{\zeta}_t S_t^{(T+h)} + \Psi(B)\left(\Lambda \boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t\right),$$

this change will affect the variance of $\{\mathbf{y}_t\}$ for $t \geq T+h$.

The filtered series is given by $a_t = \boldsymbol{\varepsilon}_t + \Pi(B)\boldsymbol{\zeta}_t S_t^{(T+h)}$. Thus the variance additive change affects $\{a_t\}$ for time $t \geq T+h$.

In practice, to model the pattern of an exogenous disturbance in a multiple time series setting, we prefer to use either (a) or (f) since the variables of the system are then allowed to incorporate the dynamics of the VAR model. Otherwise the disturbance effects will be unrealistic by affecting just one particular variable of the system.

3.3.3 The forecast

An expression that allows for a deterministic change on the forecast horizon can be obtained directly from (3.5) as

$$\mathbf{Y}_{F,D} = \mathbf{Y}_F + \mathbf{D}_F$$

where $\mathbf{Y}_{F,D} = \left(\widetilde{\mathbf{y}}'_{T+1} \cdots \widetilde{\mathbf{y}}'_{T+H}\right)'$ is a $kH \times 1$ vector containing the future observed values of the multiple time series and $\mathbf{D}_F = \left(\mathbf{d}'_{T+1} \cdots \mathbf{d}'_{T+H}\right)'$ is a $kH \times 1$ deterministic vector that accounts for the deterministic change, here $\mathbf{d}_t = \alpha(B)\boldsymbol{\omega}\xi_t^{(T+h)}$

is a $k \times 1$ vector. For example, consider the case of an innovational change, thus $\mathbf{d}_t = \Psi(B)\boldsymbol{\omega}\xi_t^{(T+h)}$, that can be written as a recursive equation

$$\Pi(B)\mathbf{d}_t = \boldsymbol{\omega}\xi_t^{(T+h)}, \text{ for } t \geq T+h \text{ and } \mathbf{d}_t = \mathbf{0} \text{ otherwise.}$$

In a similar fashion, an expression that allows for a stochastic change can be written as

$$\mathbf{Y}_{F,V} = \mathbf{Y}_F + \mathbf{V}_F$$

where $\mathbf{Y}_{F,V}$ is a $kH \times 1$ vector containing the future observed values and $\mathbf{V}_F = \left(\mathbf{v}'_{T+1} \cdots \mathbf{v}'_{T+H}\right)'$ is a $kH \times 1$ stochastic vector that accounts for a stochastic change, here $\mathbf{v}_t = \alpha(B)\boldsymbol{\zeta}_t S_t^{(T+h)}$ are $k \times 1$ random vectors. Consider for instance a variance innovational change, in this case we have $\mathbf{v}_t = \Psi(B)\boldsymbol{\zeta}_t S_t^{(T+h)}$ which can be written as

$$\Pi(B)\mathbf{v}_t = \boldsymbol{\zeta}_t S_t^{(T+h)}, \text{ for } t \geq T+h \text{ and } \mathbf{v}_t = \mathbf{0} \text{ otherwise.}$$

Thus, an expression that accounts for both deterministic and stochastic changes can be written as

$$\mathbf{Y}_{F,D,V} = \mathbf{Y}_F + \mathbf{D}_F + \mathbf{V}_F.$$

The forecast of this process, given its historical record is

$$\mathrm{E}\left(\mathbf{Y}_{F,D,V} | \, \mathbf{Y}\right) = \mathrm{E}(\mathbf{Y}_F | \mathbf{Y}) + \mathbf{D}_F,$$

and its forecast error vector can be written as

$$\begin{aligned} \mathbf{Y}_{F,D,V} - \mathrm{E}\left(\mathbf{Y}_{F,D,V} | \, \mathbf{Y}\right) &= \mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F | \mathbf{Y}) + \mathbf{V}_F \\ &= \boldsymbol{\Psi}\boldsymbol{\epsilon}_F + \mathbf{V}_F. \end{aligned}$$

So, when a variance innovational change occurs the MSE of $\mathrm{E}(\mathbf{Y}_{F,D,V}|\,\mathbf{Y})$ will take the form

$$\mathrm{MSE}\left[\mathrm{E}\left(\mathbf{Y}_{F,D,V}|\,\mathbf{Y}\right)\right] = \mathbf{\Psi}\left[\left(I \otimes \mathbf{\Sigma}_{\varepsilon}\right) + \left(\widetilde{I} \otimes \mathbf{\Sigma}_{\zeta}\right)\right]\mathbf{\Psi}'.$$

Here $\widetilde{I}$ =diag$(0\cdots01\cdots1)$ is an $H \times H$ diagonal matrix whose $h-1$ first elements are 0 and the rest are 1. The corresponding expression for a variance additive change will be

$$\mathrm{MSE}\left[\mathrm{E}\left(\mathbf{Y}_{F,D,V}|\,\mathbf{Y}\right)\right] = \mathbf{\Psi}\left(I \otimes \mathbf{\Sigma}_{\varepsilon}\right)\mathbf{\Psi}' + \left(\widetilde{I} \otimes \mathbf{\Sigma}_{\zeta}\right).$$

3.4 Restricted forecasts of a VAR with structural change

We are concerned with obtaining the vector of forecasts when additional information about the future values of the series is given in the form of linear restrictions, that is when

$$\mathbf{R} = \mathbf{C}\mathbf{Y}_{F,D,V} \tag{3.8}$$

here $\mathbf{C}$ is an $M \times kH$ matrix representing the particular linear combinations of $\mathbf{Y}_{F,D,V}$,

$$\mathbf{C} = \left(\begin{array}{ccccccc} C_{1,1} & \cdots & C_{1,k} & \cdots & C_{1,k(H-1)+1} & \cdots & C_{1,kH} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \underbrace{C_{M,1} \quad \cdots \quad C_{M,k}}_{\mathbf{y}_{T+1}} & & & \cdots & \underbrace{C_{M,k(H-1)+1} \quad \cdots \quad C_{M,kH}}_{\mathbf{y}_{T+H}} & & \end{array} \right)$$

where the rows are independent, so that the rank of $\mathbf{C}$ is M. Besides $\mathbf{R} = (r_1 \cdots r_M)'$ is the $M \times 1$ vector of values that these linear combinations will take on.

Thus the problem of finding the optimal restricted forecast vector (in MSE sense) of $\mathbf{Y}_{F,D,V}$ can be posed as the Lagrangian minimization of

$$\begin{aligned} L &= \left[\mathbf{Y}_{F,D,V} - \mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right)\right]' \boldsymbol{\Sigma}_{F,D,V}^{-1} \left[\mathbf{Y}_{F,D,V} - \mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right)\right] \\ &\quad + 2\boldsymbol{\lambda}' \left(\mathbf{R} - \mathbf{C}\mathbf{Y}_{F,D,V}\right) \end{aligned}$$

where $\boldsymbol{\Sigma}_{F,D,V} \equiv$MSE$[\mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right)]$. Thus, $\partial L / \partial \mathbf{Y}_{F,D,V} = 0$ implies

$$\widehat{\mathbf{Y}}_{F,D,V} = \mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right) + \boldsymbol{\Sigma}_{F,D,V}\mathbf{C}'\widehat{\boldsymbol{\lambda}}$$

and $\partial L / \partial \boldsymbol{\lambda} = 0$ implies

$$\mathbf{R} = \mathbf{C}\widehat{\mathbf{Y}}_{F,D,V}.$$

These last two equations lead us to

$$\widehat{\boldsymbol{\lambda}} = \left(\mathbf{C}\boldsymbol{\Sigma}_{F,D,V}\mathbf{C}'\right)^{-1} \left[\mathbf{R} - \mathbf{C}\mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right)\right],$$

thus the restricted forecast of a VAR process with structural change in the forecast horizon is

$$\widehat{\mathbf{Y}}_{F,D,V} = \mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right) + \mathbf{A}\left[\mathbf{R} - \mathbf{C}\mathrm{E}\left(\mathbf{Y}_{F,D,V} \mid \mathbf{Y}\right)\right]$$

where $\mathbf{A} = \boldsymbol{\Sigma}_{F,D,V}\mathbf{C}'\left(\mathbf{C}\boldsymbol{\Sigma}_{F,D,V}\mathbf{C}'\right)^{-1}$. The mean squared error of this vector of forecasts is

$$\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,D,V}\right) = \left(I - \mathbf{AC}\right)\boldsymbol{\Sigma}_{F,D,V}.$$

3.5 Empirical illustration

As already mentioned, several economic reforms (fiscal, energetic, pensions, etc.) have been proposed as necessary to achieve better growth in the Mexican economy. The lack of such reforms has been considered the main reason for the stagnation of the economy. Since an economic reform in the short-run seems likely to occur, we consider the Mexican case appropriate to illustrate our methodology. To that end let us pretend that a reform will occur in year 2005 and that this will produce a recovering effect on the economy. We estimate a six-dimensional integrated VAR model with Mexican data and get restricted forecasts that fulfill the 2004 government targets and account for the 2005 expected structural change. Two scenarios are considered: one in which the structural change is presented as a deterministic change and other in which the change is considered to be stochastic. Program routines for estimation and forecasting of the system were done in Matlab 6.5-Release 13 (MathWorks, Inc. Software).

3.5.1 The data

When the Mexican Government announced the targets for year 2004 the available data ran up to 2003:III. The data set consists of 59 quarterly observations covering a period from 1989:I to 2003:III. A description of the variables that conform the system is as follows.

Mexican monthly inflation rate (PMEX): First difference of log Consumer Price Index (ipcmex) with base 1994=100. Source: *Bank of Mexico*. The Consumer Price Index is a monthly series and the quarterly series is obtained with the values at the end of each quarter. $\text{PMEX}_t = \ln(\text{ipcmex}_t) - \ln\left(\text{ipcmex}_{t-1}\right)$.

Gross domestic product (LGDP): Measured in thousands of Mexican pesos at constant prices of 1993. Source: *Instituto Nacional de Estadística, Geografía e Informática, National Accounts System*. The series is log transformed. $\text{LGDP}_t = \ln(\text{GDP}_t)$.

Real demand of money (LMONB): Currency held by the public plus domestic currency and checking accounts in resident banks. This is a monthly series given in nominal terms in thousand of Mexican pesos (MONB). The quarterly series is obtained by averaging the monthly values and is deflated by ipcmex. Source: *Bank of Mexico.* The series is log transformed. $\text{LMONB}_t = \ln(\text{MONB}_t/\text{ipcmex}_t)$.

Trade balance deficit (TRDB): Defined as income minus expenditure of the foreign sector. This is a quarterly series given in millions of dollars (DEF). Source: *Bank of Mexico.* The series is transformed by dividing it by 10,000 to homogenize the data scales. $\text{TRDB}_t = \text{DEF}_t/10{,}000$.

Unemployment rate (LUNMP): Source: *Instituto Nacional de Estadística Geografía e Informática, National Urban Employment Survey.* The unemployment data (UNMP) are log transformed. $\text{LUNMP}_t = \ln(\text{UNMP}_t)$.

US monthly inflation rate (PUSA): First difference of log US Consumer Price Index (ipcusa) with base 1982-84=100. Source: *US Department of La-*

bor, Bureau of Labor Statistics. The US Consumer Price Index is monthly and the quarterly observations are the values at the end of each quarter. $\mathrm{PUSA}_t = \ln(\mathrm{ipcusa}_t) - \ln(\mathrm{ipcusa}_{t-1})$.

We consider PMEX, GDP, MONB, DEF, UNMP and PUSA as the original variables. The series were transformed basically to stabilize their variances. Thus, in this case the vector of variables that will be used is $\mathbf{y}_t = F(\mathbf{x}_t)$ where $\mathbf{x}_t$ is the $k \times 1$ vector of original variables at time t and $F(\cdot)$ is an appropriate transformation.

3.5.2 Order of integration

The order of integration of the transformed series was decided on the basis of Augmented Dickey-Fuller (ADF) tests. The regression model includes a constant term, centered dummies for all the variables and a deterministic trend for LGDP_t and LMONB_t. The general equation is

$$\Delta z_t = a + b_0 t + \sum_{i=1}^{n} b_i \delta_{it} + c_0 z_{t-1} + \sum_{j=1}^{p} c_j \Delta z_{t-j} + \varepsilon_t$$

In order to account for the Mexican crisis of year 1995 we included in the analysis two dummy variables for the first and second quarters of that year. Table 3.1 shows the results of the ADF tests with and without the dummy variables. That is, with and without accounting for structural change (SC).

The τ statistic in Table 3.1 allows one to test $H_0 : c_0 = 0$. The symbol (*) indicates rejection of the null hypothesis at the 5% significance level. The order,

p, of the model was selected to guarantee no residual autocorrelation. Critical values do not consider intervention variables to account for the 1995 crisis. Except for domestic inflation (PMEX) in levels, all variables are $I(1)$ and the order of integration does not depend on the inclusion of dummy variables, but there still is some doubt whether or not inflation is stationary. However, since the result can be distorted by the inclusion of dummy variables, we assume here that PMEX is $I(1)$.

Table 3.1: ADF unit root tets results

Variable	$H_0 : I(1)$				$H_0 : I(2)$			
	Without SC		With SC		Without SC		With SC	
	p	τ	p	τ	p	τ	p	τ
PMEX	2	−1.93	2	−3.93*	1	−8.14*	0	−9.65*
LGDP	0	−2.20	1	−2.65	0	−6.72*	0	−10.26*
LMONB	1	−1.60	4	−0.46	0	−5.01*	2	−6.92*
TRDB	0	−2.24	4	−1.78	0	−6.96*	3	−4.44*
LUNMP	1	−1.56	1	−2.53	0	−5.47*	0	−6.84*
PUSA	2	−2.65	–	–	1	−9.68*	–	–

The symbol * indicates rejection of H_0 at the 5% significance level.

3.5.3 VEC estimation

The VEC model includes the six endogenous economic variables described previously, $\mathbf{y}_t = (\text{PMEX}_t, \text{LGDP}_t, \text{LMONB}_t, \text{TRDB}_t, \text{LUNMP}_t, \text{PUSA}_t)'$, a constant term, centered dummy variables to account for seasonal effects and two dummy variables to account for the 1995 economic crisis, that is $\mathbf{D}_t = (\text{const}, S_{1,t}, S_{2,t}, S_{3,t}, I_{95:I,t}, I'_{95:II,t})$. Although the USA inflation is basically exogenous to the Mexican economy, it was considered as an endogenous variable since not other variable in the system affects PUSA significantly. The system turned out to be CI$(1, 1)$ and an integrated VAR(3) model provided a reasonable representation for the variables in levels. Therefore the system was estimated with the following VEC model

$$\Delta\mathbf{y}_t = \Lambda\mathbf{D}_t + \boldsymbol{\gamma}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \Pi_1^*\Delta\mathbf{y}_{t-1} + \Pi_2^*\Delta\mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t$$

The estimation results are as follows (t-values in parentheses)

$$\widehat{\boldsymbol{\gamma}} = \begin{pmatrix} \underset{(-1.75)}{-0.289} & \underset{(3.80)}{0.746} & \underset{(-1.16)}{-0.451} & \underset{(-1.47)}{-1.926} & \underset{(-1.50)}{-2.362} & \underset{(0.65)}{0.047} \end{pmatrix},$$

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} 1 & -0.029 & 0.043 & -0.045 & 0.01 & -0.446 \end{pmatrix},$$

$$\widehat{\Lambda} = \begin{pmatrix} - & - & \underset{(-2.32)}{-0.029} & - & \underset{(9.19)}{0.112} & \underset{(5.37)}{0.100} \\ - & \underset{(-4.03)}{-0.078} & \underset{(-4.04)}{-0.061} & \underset{(-7.74)}{-0.090} & \underset{(-4.42)}{-0.064} & \underset{(-3.94)}{-0.087} \\ \underset{(2.66)}{0.020} & \underset{(-2.62)}{-0.101} & \underset{(-5.25)}{-0.156} & \underset{(-5.93)}{-0.136} & \underset{(-5.02)}{-0.143} & \underset{(-2.92)}{-0.128} \\ - & \underset{(3.22)}{0.418} & \underset{(2.42)}{0.243} & - & \underset{(6.38)}{0.615} & \underset{(2.67)}{0.393} \\ - & - & \underset{(2.01)}{0.243} & - & \underset{(2.65)}{0.307} & - \\ - & \underset{(2.71)}{0.019} & - & \underset{(2.35)}{0.010} & - & - \end{pmatrix},$$

$$\widehat{\Pi}_1^* = \begin{pmatrix} \underset{(-2.04)}{-0.327} & - & - & - & - & - \\ \underset{(-3.45)}{-0.658} & \underset{(-3.35)}{-0.468} & \underset{(2.93)}{0.251} & \underset{(2.47)}{0.056} & - & - \\ - & - & - & - & - & - \\ - & - & \underset{(-2.06)}{-1.173} & \underset{(-3.80)}{-0.580} & \underset{(2.15)}{0.312} & \underset{(2.18)}{6.959} \\ \underset{(2.08)}{3.181} & - & - & - & - & - \\ - & - & - & - & - & \underset{(-3.16)}{-0.550} \end{pmatrix},$$

$$\widehat{\Pi}_2^* = \begin{pmatrix} \underset{(-2.36)}{-0.225} & - & - & - & - & \underset{(2.07)}{0.757} \\ \underset{(-2.46)}{-0.279} & - & \underset{(2.31)}{0.216} & - & - & - \\ - & - & - & - & - & - \\ - & - & - & \underset{(-2.56)}{-0.325} & - & \underset{(3.13)}{9.085} \\ - & - & - & - & - & - \\ - & - & - & - & - & - \end{pmatrix},$$

with R^2 equal to 0.83, 0.93, 0.88, 0.67, 0.52 and 0.70 for ΔPMEX, ΔLGDP, ΔLMONB, ΔTRDB, ΔLUNMP and ΔPUSA, respectively. The matrices $\widehat{\Lambda}$, $\widehat{\Pi}_1^*$, and $\widehat{\Pi}_2^*$ show only the numerical values of those elements found significant at the 5% level.

The following matrix shows the contemporaneous residual correlations.

$$
\begin{array}{c}
\\ \text{PMEX} \\ \text{LGDP} \\ \text{LMONB} \\ \text{TRDB} \\ \text{LUNMP} \\ \text{PUSA}
\end{array}
\begin{array}{c}
\begin{array}{cccccc} \text{PMEX} & \text{LGDP} & \text{LMONB} & \text{TRDB} & \text{LUNMP} & \text{PUSA} \end{array} \\
\begin{bmatrix}
1.00 & & & & & \\
0.24 & 1.00 & & & & \\
-0.42 & 0.02 & 1.00 & & & \\
-0.10 & -0.02 & 0.25 & 1.00 & & \\
-0.30 & -0.14 & -0.01 & -0.08 & 1.00 & \\
0.28 & 0.16 & -0.21 & 0.32 & -0.22 & 1.00
\end{bmatrix}
\end{array}
$$

Here we can appreciate the correlation between Mexican inflation rate and real demand for money, with negative sign as expected. It is also worth noting the positive correlation between USA inflation rate and trade balance deficit. In Table 3.2 we report the results of Johansen tests.

Table 3.2: Johansen cointegration analysis

Null	Trace Statistic	Crit 95%	Crit 99%	Eigen Statistic	Crit 95%	Crit 99%
$r \leq 0$	131.10**	95.75	104.96	54.43**	40.08	45.87
$r \leq 1$	76.68*	69.82	77.82	36.29*	33.88	39.37
$r \leq 2$	40.38	47.86	54.68	21.47	27.59	32.72
$r \leq 3$	18.91	29.80	35.46	13.74	21.13	25.87
$r \leq 4$	5.17	15.49	19.94	4.63	14.26	18.52
$r \leq 5$	0.54	3.84	6.64	0.54	3.84	6.64

The symbol ** indicates rejection at 1% level and * rejection at 5%.

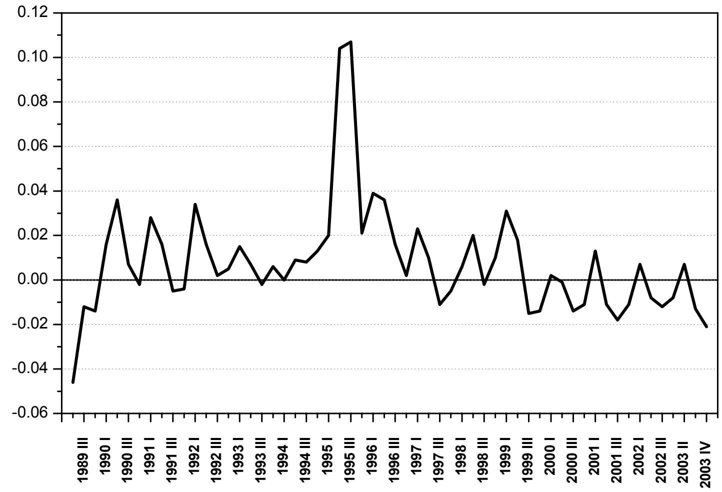


Figure 3.1: Cointegration relationship

At the 5% significance level there are two cointegrating relationships and only one at the 1% level. We decided to use only one cointegration relationship. The resulting cointegration equation is

$$\epsilon_t = \mathrm{PMEX}_t - 0.029\mathrm{LGDP}_t + 0.043\mathrm{LMONB}_t - 0.045\mathrm{TRDB}_t + 0.011\mathrm{LUNMP}_t - 0.446\mathrm{PUSA}_t.$$

whose graph is shown in Fig. 3.1. The observed and fitted series of the VEC model as well as their residuals are shown in Fig. 3.2.

3.5.4 Restricted forecasts of the Mexican economy

At the end of 2003 the Mexican Government published the economic policies as well as the economic targets for 2004 (see SHCP, 2003). There, it was foreseen that the rate of growth of GDP would move from 1.5% in 2003 to 3.1% in 2004.

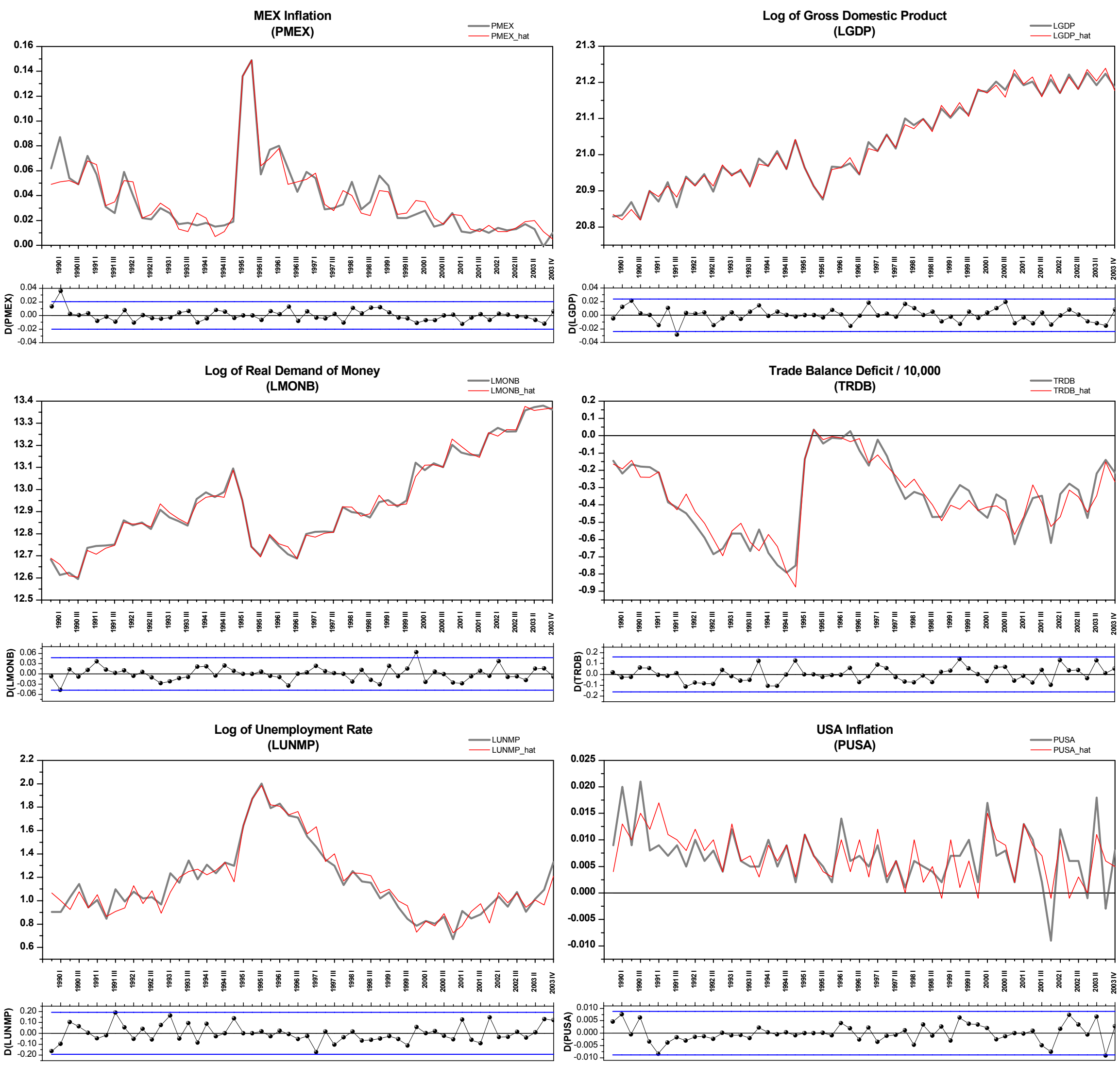


Figure 3.2: Observed and estimated series in levels and corresponding residuals of their first differences, with $\pm 2\sigma$ bands plotted as horizontal lines.

The annual inflation rate was targeted to be reduced from 3.8% in 2003 to 3.0% in 2004 and the trade balance deficit was supposed to move from -11,847.0 (-1.9% of GDP in 2003) to -16,419.9 (-2.6% of GDP target in 2004). That is

$$\mathrm{PMEX}_{2004:I} + \mathrm{PMEX}_{2004:II} + \mathrm{PMEX}_{2004:III} + \mathrm{PMEX}_{2004:IV} = \ln(1.03),$$

$$\mathrm{LGDP}_{2004:IV} - \mathrm{LGDP}_{2003:IV} = \ln(1.031),$$

$$\mathrm{TRDB}_{2004:I} + \mathrm{TRDB}_{2004:II} + \mathrm{TRDB}_{2004:III} + \mathrm{TRDB}_{2004:IV} = -16,419.9/10,000.$$

Additionally, we suppose that the economic reform will have a positive effect in the economy leading the inflation rate (PMEX) down to 3% and the annual rate of growth of GDP up to 7%, so that

$$\mathrm{PMEX}_{2005:I} + \mathrm{PMEX}_{2005:II} + \mathrm{PMEX}_{2005:III} + \mathrm{PMEX}_{2005:IV} = \ln(1.03),$$

$$\mathrm{LGDP}_{2005:IV} - \mathrm{LGDP}_{2004:IV} = \ln(1.07).$$

Thus, the linear restriction (3.8) considers an $H = 9$ periods ahead forecast (to reach the end of 2005) and gets specified by

$$\mathbf{R} = \begin{pmatrix} \ln(1.03) \\ \ln(1.031) \\ -16,419.9/10,000 \\ \ln(1.03) \\ \ln(1.07) \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} \mathbf{0} & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{e}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{e}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{e}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 & \mathbf{e}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{e}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{e}_2 \end{pmatrix}$$

where $\mathbf{0} = (0 \cdots 0)$ is the 1×6 zero matrix and $\mathbf{e}_i = (0 \cdots 1 \cdots 0)$ is a 1×6 vector with 1 in the ith position and zeros elsewhere.

Notice that $\mathbf{R}$ can be rewritten as $\mathbf{R} = (\mathbf{R}_1'|\mathbf{R}_2')'$ where $\mathbf{R}_1$ is an $M_1 \times 1$ vector that collects the government economic targets in 2004 (with $M_1 = 3$) and $\mathbf{R}_2$ is an $M_2 \times 1$ vector that collects the expected structural change in 2005 (with $M_2 = 2$). Similarly, the matrix $\mathbf{C} = (\mathbf{C}_1'|\mathbf{C}_2')'$ can be partitioned by rows. We use this partition in the stochastic change scenario shown below.

3.5.4.1 Scenario 1, deterministic innovational change

We assume that the expected economic reform will affect primarily the LGDP and PMEX variables and that PUSA is an exogenous variable. The rest of the variables will be affected by the economic reform through a deterministic innovational change that takes into account the system dynamics.

So, let the time path for LGDP and PMEX be determined by a gradual change, i.e.

$$\Delta\,(1 - \rho B)\, d_{it} = \iota_i \xi_t^{(T+h)}, \text{ with } 0 < \rho < 1. \tag{3.9}$$

while $d_{it} = 0$ for PUSA. Here, d_{it} denotes the ith element of $\mathbf{d}_t$ and ι_i is the impact disturbance variable on the ith variable.

To compute the impact disturbance variable ι_i from these paths we consider that $\mathrm{E}(\mathbf{Y}_{F,D}|\mathbf{Y})$ achieves exactly the expected structural change for both LGDP and PMEX. Therefore, $\mathbf{D}_F$ should satisfy the corresponding restrictions

$$C_i\mathbf{D}_F = r_i - C_iE(\mathbf{Y}_F|\mathbf{Y}) \tag{3.10}$$

where C_i is the ith row of $\mathbf{C}$.

For the LGDP variable, equation (3.10) becomes $d_{i,T+h+N} - d_{i,T+h-1} = r_i - C_iE(\mathbf{Y}_F|\mathbf{Y})$ while from (3.9), $d_{i,T+h+N} = \iota_i\left(1+\rho+\cdots+\rho^N\right)$ and $d_{i,T+h-1} = 0$. Thus, $\iota_i = [r_i - C_i\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})] / \left(1+\rho+\cdots+\rho^N\right)$. Here $N+1$ is the number of periods since the change takes place until the desired level is achieved. Proceeding in the same way for PMEX, equation (3.10) takes the form $d_{i,T+h} + \cdots + d_{i,T+h+N} = r_i - C_iE(\mathbf{Y}_F|\mathbf{Y})$ and $d_{i,T+h+n} = \iota_i\left(1+\rho+\cdots+\rho^n\right)$ for $n = 1,\ldots,N$. Thus, the impact disturbance variable for PMEX turns out to be $\iota_i = [r_i - C_i\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})] / \left[(N+1)+N\rho+\cdots+\rho^N\right]$. We chose the value $\rho = 0.6$ as suggested by Tsay (1988).

We make use of the following multiple disturbance innovational change model

$$\Pi(B)\mathbf{d}_t = \sum_{n=0}^{N}\boldsymbol{\omega}_t\xi_t^{(T+h+n)}, \text{ for } t \geq T+h \text{ and } \mathbf{d}_t = \mathbf{0} \text{ otherwise.} \tag{3.11}$$

Here the impact vectors $\boldsymbol{\omega}_t$ for $t = 1,\ldots,T+H$ were computed recursively using the above deterministic change on LGDP, PMEX and PUSA variables. That is,

$\boldsymbol{\omega}_t = \mathbf{0}$ for $t < T + h$ and the ith entry of $\boldsymbol{\omega}_{T+h+n}$ for $n \geq 0$ corresponding to LGDP or PMEX is given by

$$\omega_{i,T+h+n} = d_{i,T+h+n} - (\Pi_1 \mathbf{d}_{T+h+n-1} + \cdots + \Pi_p \mathbf{d}_{T+h+n-p})_i \text{ for } n = 0, ..., N$$

and that for PUSA is given by

$$\omega_{i,T+h+n} = -\left(\Pi_1 \mathbf{d}_{T+h+n-1} + \cdots + \Pi_p \mathbf{d}_{T+h+n-p}\right)_i ,$$

the rest of the elements of $\boldsymbol{\omega}_{T+h+n}$ were set equal to zero because no restrictions apply to their corresponding variables. Putting this into (3.11) gives $\mathbf{d}_{T+h+n}$, then we can compute $\boldsymbol{\omega}_{T+h+n+1}$ in the same manner and so on.

Fig. 3.3 shows the results of the deterministic change $\mathbf{D}_F$ on the system. Fig. 3.4 shows the unrestricted and restricted VAR forecasts with deterministic change along with their 90% probability intervals for the six original economic variables. Table 3.3 reports the forecasts for the original variables. The unrestricted and restricted forecast of the original variables are denoted by $\widehat{\mathbf{X}}_F = F^{-1}\left(\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})\right)$ and $\widehat{\mathbf{X}}_{F,D} = F^{-1}\left(\widehat{\mathbf{Y}}_{F,D}\right)$ respectively, where $F^{-1}(\cdot)$ is the inverse of the transformation $F(\cdot)$ employed before. The Mexican Government economic targets for year 2004, the assumed deterministic change in year 2005 and the unrestricted forecast with deterministic change, $\widehat{\mathbf{X}}_F + \mathbf{D}_F^* = F^{-1}\left(\mathrm{E}\left(\mathbf{Y}_{F,D}|\,\mathbf{Y}\right)\right)$ are also reported in that table. By construction, the restricted forecasts of the process with deterministic change, $\widehat{\mathbf{X}}_{F,D}$, attain the economic targets for 2004 as well as the expected deterministic change in 2005 exactly.

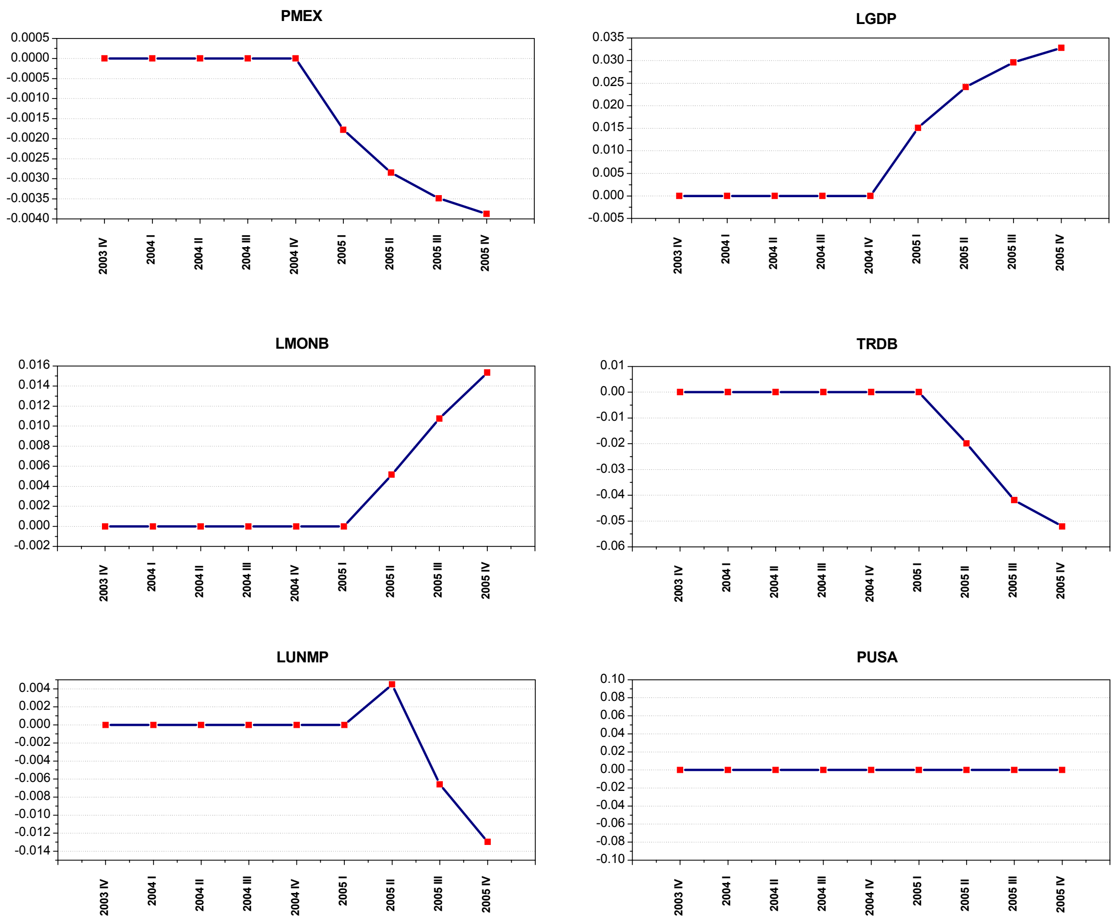


Figure 3.3: Innovational deterministic change for the system. (Transformed) Mexican prices and GDP follow a gradual change to achieve their corresponding targets for year 2005, PUSA is unaffected by these changes.

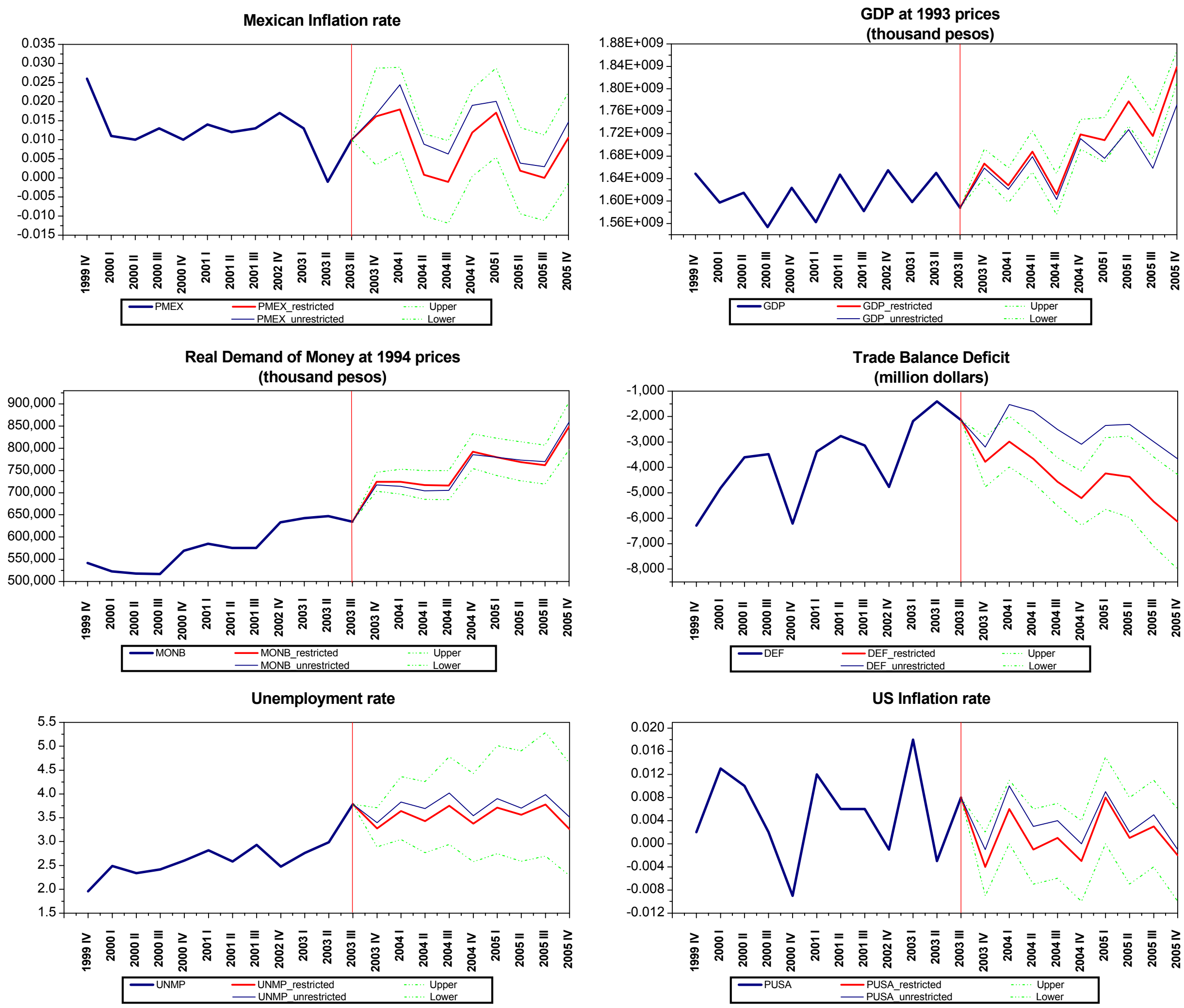


Figure 3.4: Unrestricted and restricted forecasts with a deterministic structural change and 90% probability intervals (forecast origin at 2003:III).

Table 3.3: VEC forecasts results for the Mexican system

Variable	2004			2005			
	Gov. Target	$\widehat{X}_F$	$\widehat{X}_{F,D}$	Det. Change	$\widehat{X}_F$	$\widehat{X}_F + D_F^*$	$\widehat{X}_{F,D}$
PMEX	3.00%	6.03%	3.00%	3.00%	4.24%	3.00%	3.00%
GDP	3.10%	3.16%	3.10%	7.00%	3.54%	7.00%	7.00%
MONB	–	9.51%	9.41%	–	9.38%	11.07%	7.13%
DEF	−16419.9	−8923.2	−16419.9	–	−11325.1	−12464.6	−20096.8
UNMP	–	3.54%	3.38%	–	3.51%	3.47%	3.26%
PUSA	–	1.72%	0.26%	–	1.61%	1.61%	1.00%

It is worth emphasizing the fact that the unrestricted and restricted forecasts are drastically different. For instance, the unrestricted forecast of the Mexican inflation rate (PMEX) in year 2004 is 6.03%, twice the government target, and will still be controlled in year 2005 around 3.00% if an economic reform occurs. The restricted forecast of the rate of growth of GDP in year 2004 is near to the unrestricted one, however, an economic reform in year 2005 will increase significantly the rate of growth of GDP. Except for MONB, in year 2005 the unrestricted forecasts with structural change added, $\widehat{\mathbf{X}}_F + \mathbf{D}_F^*$, lie between the unrestricted and the restricted forecasts. This behavior of MONB can be explained by the significant correlation between TRDB and LMONB. Another interesting point to note in this system is that if the restricted forecast for PMEX increase to its unrestricted path then the real demand of money will decrease as shown in Fig. 3.5.

3.5.4.2 Scenario 2, variance innovational change

When the stochastic structure is deemed to change due to the economic reform in year 2005 the contaminating covariance matrix can be used to introduce uncertainty. As in the deterministic change, we first introduce the economic reform effect to PMEX and LGDP and propagate it to the rest of the variables by assuming a variance innovational change. Consider for example that the contaminating covariance matrix is given by $\mathbf{\Sigma}_{\boldsymbol{\zeta}} = a\mathbf{Q}$, here the contaminating parameter is $a > 0$ and $\mathbf{Q} = \mathrm{diag}(r_4^2, r_5^2, 0, \ldots, 0)$ is a 6×6 matrix. Notice that the contaminating effect appears in PMEX and LGDP and is proportional to r_4^2 and to r_5^2 respectively. The problem now lies in selecting the contaminating parameter a appropriately. To that end, let us define the following distance vector

$$\boldsymbol{\eta} \equiv \mathbf{R}_2 - \mathbf{C}_2 \mathrm{E}\left(\mathbf{Y}_{F,V} \middle| \mathbf{Y}\right) = \mathbf{C}_2 \mathbf{\Psi} \boldsymbol{\epsilon}_F + \mathbf{C}_2 \mathbf{V}_F$$

then $\mathbf{R}_2$ and $\mathrm{E}(\mathbf{Y}_{F,V} | \mathbf{Y})$ are said to be compatible if the distance vector is close to zero.

From the normality assumption of $\boldsymbol{\epsilon}_F$ and $\boldsymbol{\zeta}_t^{T+h}$ we have that

$$\boldsymbol{\eta} \sim N\left(\mathbf{0}, \mathbf{\Omega}(a)\right)$$

where $\mathbf{\Omega}(a) = \mathbf{C}_2 \mathbf{\Psi} \left[(I \otimes \mathbf{\Sigma}_\varepsilon) + a \left(\widetilde{I} \otimes \mathbf{Q} \right) \right] \mathbf{\Psi}' \mathbf{C}_2'$ is an $M_2 \times M_2$ symmetric positive semidefinite covariance matrix. So, if $\mathbf{\Omega}(a)$ is non-singular, a statistic for

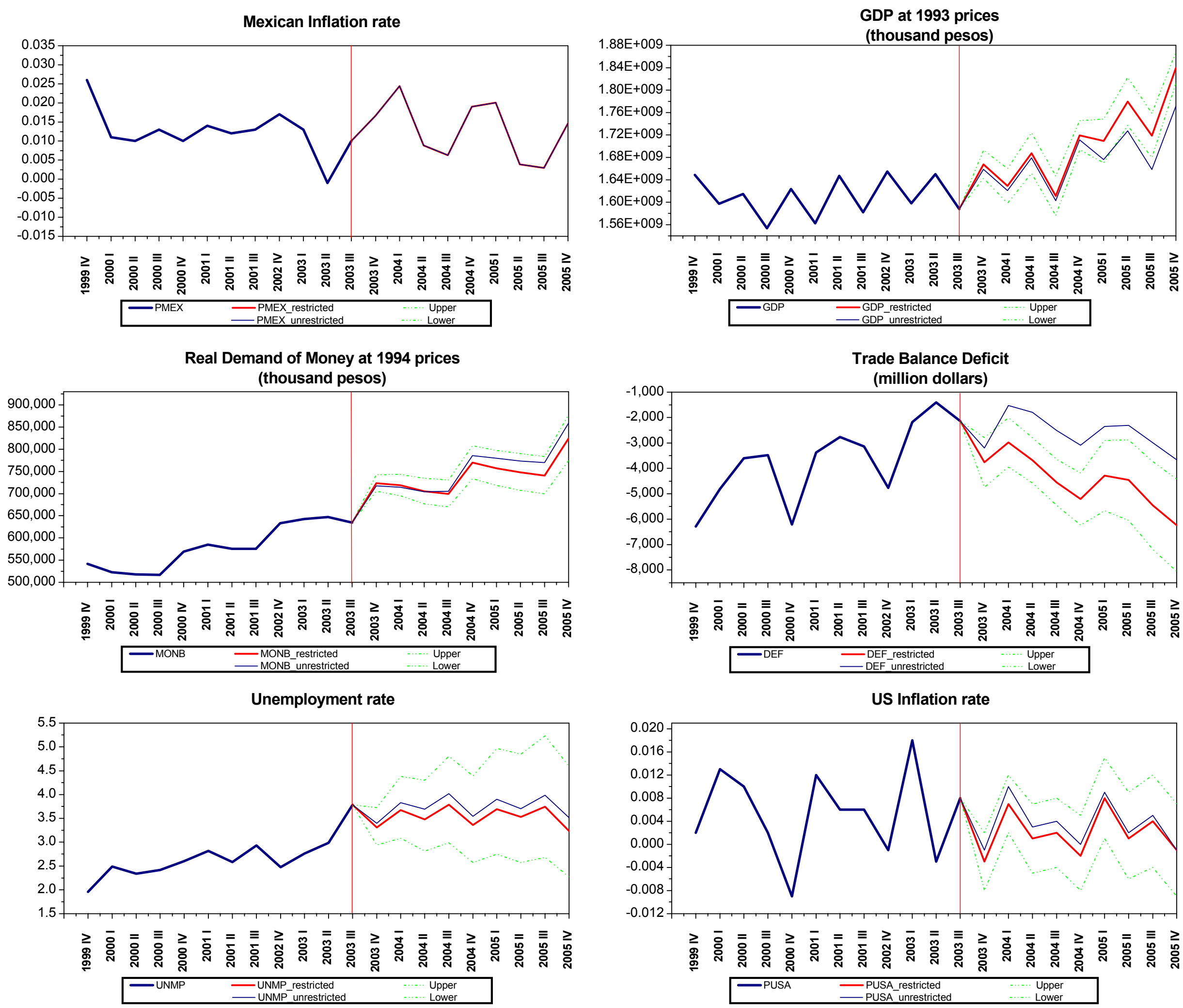


Figure 3.5: Unrestricted and restricted forecasts with a deterministic structural change and 90% probability intervals. The restricted forecast of PMEX follows its unrestricted path (forecast origin at 2003:III).

testing compatibility between restrictions and unrestricted forecasts is defined as

$$\mathbf{K}(a) \equiv \boldsymbol{\eta}'\boldsymbol{\Omega}^{-1}(a)\boldsymbol{\eta} \sim \chi^2_{M_2},$$

which is of the type derived in Guerrero *et al.* (2005). We say that $\boldsymbol{\eta}$ is in the compatibility region at a level α if

$$\mathbf{K}(a) \leq \chi^2_{M_2}(\alpha)$$

where $\chi^2_{M_2}(\alpha)$ denotes the upper α percentage point of the $\chi^2_{M_2}$ distribution. Since $\mathbf{K}(a)$ tends to zero as the contaminating parameter a goes to infinity (see the Appendix) we should choose a large enough to obtain compatibility, but keeping in mind that the larger the value of a the greater the uncertainty in the forecasts.

In the present case the compatibility test does not reject the compatibility hypothesis for the typical significance levels. Nonetheless, just for illustrative purposes we decided to take $\alpha = 0.5$ so that $\chi^2_2(0.5) = 1.386$. Thus, a contaminating parameter $a = 0.040$ was required. Fig. 3.6 shows the unrestricted and restricted forecasts with stochastic change of the original variables as well as 90% probability intervals. Table 3.4 contains the restricted forecasts with stochastic change of the original variables, $\widehat{\mathbf{X}}_{F,V} = F^{-1}\left(\widehat{\mathbf{Y}}_{F,V}\right)$.

As pointed out above, consideration of extra model information helps the forecasts to not only depend on the dynamics contained in the historical record but to take into account what the government wants to attain. Except for MONB and PUSA, the figures are similar to those shown in Table 3.3 for $\widehat{\mathbf{X}}_{F,D}$. The main

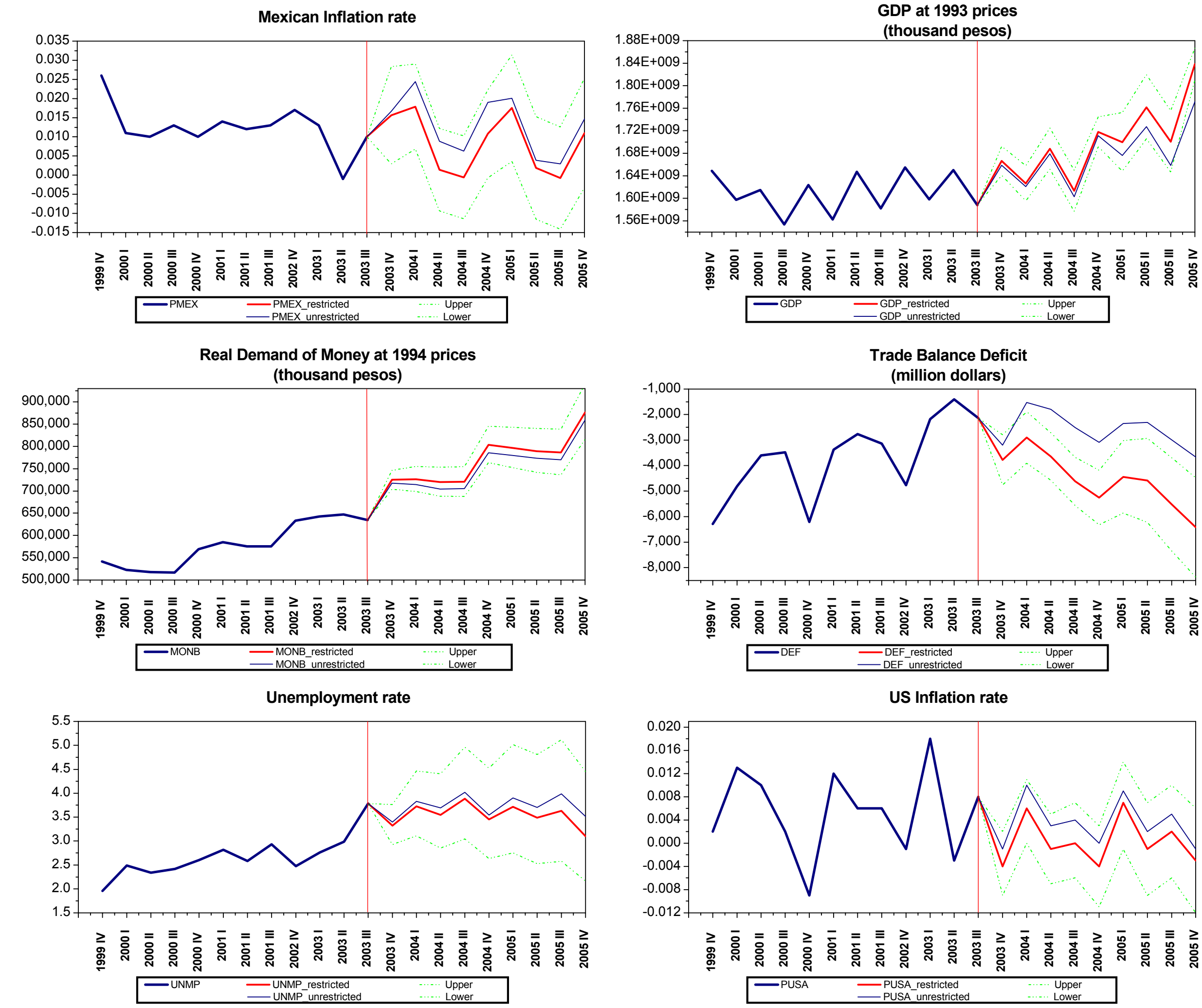


Figure 3.6: Unrestricted and restricted forecasts of the VAR with a variance innovational change and 90% probability intervals (forecast origin at 2003:III).

difference with the previous scenario is that uncertainty induces rates of growth of MONB higher than those reported for the deterministic change.

Table 3.4: Restricted forecasts with stochastic change for the Mexican system, $\widehat{\mathbf{X}}_{F,V}$

Variable	2004	2005
PMEX	3.00%	3.00%
GDP	3.10%	7.00%
MONB	10.84%	9.02%
DEF	−16419.9	−20952.9
UNMP	3.45%	3.10%
PUSA	0.07%	0.50%

3.6 Conclusions

This paper presents some extensions of the multivariate restricted forecasting methodology which allows one to take into account an expected structural change in the forecast horizon. The types of structural changes are presented in a multivariate setting as a generalization of those introduced by Tsay (1988). Since the forecasts should take into account the dynamics of the model we prefer to use either the deterministic innovational change or the variance innovational change. A combination of those changes can also be entertained, but the limited amount of information provided by the restrictions available, as well as confusion of deterministic and stochastic effects prevented us from carrying out the exercise.

The restricted forecast formulas of a VAR with structural change are derived by Lagrangian optimization. The methodology was illustrated with a six-dimensional system for the Mexican economy, where an economic reform has been pushed through by economic and political sectors. The application assumes that the economic targets for 2004 as well as the expected effect of an economic reform for 2005 will be reached with certainty. It is also assumed that the economic reform will modify either the deterministic or the stochastic part of a VAR model and that its effect will initially impact GDP and prices. In the empirical illustration here presented, we obtained basically the same results with either the deterministic innovational change or the variance innovational change. Thus, in this case, the choice depends on the simplicity of application and, in that sense, it is easier to apply the variance innovational change formulation, but the interpretation of effects is less clear with this approach than with the deterministic one. The combination of extra model information with VAR forecasts results in more realistic predictions since, in our case, the policies that the government has in mind are taken into account.

3.7 Appendix. Shape of $\mathbf{K}(a)$

Here we explore the dependence of the compatibility test statistic $\mathbf{K}(a)$ on the contaminating parameter a. By construction the compatibility test is non-negative. Now, if $\mathbf{C}_2\mathbf{\Psi}\left(\widetilde{I}\otimes\mathbf{Q}\right)\mathbf{\Psi}'\mathbf{C}_2'$ is a symmetric positive definite matrix (then non-singular) we have that

$$\begin{aligned}\lim_{a\to\infty}\mathbf{K}(a) &= \boldsymbol{\eta}'\left[\lim_{a\to\infty}a^{-1}\mathbf{C}_2\mathbf{\Psi}\left(I\otimes\mathbf{\Sigma}_\varepsilon\right)\mathbf{\Psi}'\mathbf{C}_2' + \mathbf{C}_2\mathbf{\Psi}\left(\widetilde{I}\otimes\mathbf{Q}\right)\mathbf{\Psi}'\mathbf{C}_2'\right]^{-1}\boldsymbol{\eta}\lim_{a\to\infty}a^{-1} \\ &= 0\end{aligned}$$

which means that compatibility can be reached by increasing the contamination of the process.

Furthermore, $\mathbf{K}(a)$ is a non-increasing function. To see that let us first compute its differential

$$d\mathbf{K} = -\boldsymbol{\eta}'\mathbf{\Omega}^{-1}\left(d\mathbf{\Omega}\right)\mathbf{\Omega}^{-1}\boldsymbol{\eta}$$

where

$$d\mathbf{\Omega} = \mathbf{C}_2\mathbf{\Psi}\left(\widetilde{I}\otimes\mathbf{Q}\right)\mathbf{\Psi}'\mathbf{C}_2'da,$$

see Magnus and Neudecker, (2002). So, the derivative of $\mathbf{K}(a)$ is equal to

$$\frac{d\mathbf{K}}{da} = -\boldsymbol{\eta}'\mathbf{\Omega}^{-1}\left[\mathbf{C}_2\mathbf{\Psi}\left(\widetilde{I}\otimes\mathbf{Q}\right)\mathbf{\Psi}'\mathbf{C}_2'\right]\mathbf{\Omega}^{-1}\boldsymbol{\eta}$$

which is non-positive. Thus, $\mathbf{K}(a)$ is a non-negative non-increasing function that tends to zero as the parameter a tends to infinity.

BIBLIOGRAPHY

[1] Box GEP, Tiao GC. 1975. Intervention analysis with applications to environmental and economic problems. *Journal of the American Statistical Association* **70**: 70-79.

[2] Engle RF, Granger CWJ. 1987. Co-integration and error correction representation, estimation and testing. *Econometrica* **55**: 251-276.

[3] Guerrero VM. 1989. Optimal conditional ARIMA forecast. *Journal of Forecasting* **8**: 215-229.

[4] Guerrero VM. 1991. ARIMA forecast with restrictions derived from a structural change. *International Journal of Forecasting* **7**: 339-347.

[5] Guerrero VM, Pena B, Senra E, Alegría A. 2005. Restricted forecasting with a VEC model: Validating the feasibility of economic targets. *Available in Eva Senra's website http://www2.uah.es/* .

[6] Lucas RE. 1976. Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* **1**: 19-46.

[7] Magnus JR, Neudecker H. 2002. *Matrix Differential Calculus*, Revised edn. John Wiley & Sons.

[8] Pandher GS. 2002. Forecasting Multivariate Time Series with Linear Restric-

tions using Constrained Structural State-space Models. *Journal of Forecasting* **21**: 281-300.

[9] Pankratz A. 1989. Time series forecasts and extra-model information. *Journal of Forecasting* **8**: 75-83.

[10] Park JY, Phillips PCB. 1988. Statistical Inference in Regression with Integrated Processes: Part 1. *Econometric Theory* **4**: 468 - 497.

[11] Park JY, Phillips PCB. 1989. Statistical Inference in Regression with Integrated Processes: Part 2. *Econometric Theory* **5**: 95 - 131.

[12] SHCP. 2003. *Criterios Generales de Política Económica para la iniciativa de ley de ingresos y el proyecto de presupuesto de egresos de la federación correspondiente al año 2004*. México: Secretaría de Hacienda y Crédito Público.

[13] Tsay RS. 1988. Outliers, level shifts, and variance changes in time series. *Journal of Forecasting* **7**: 1 - 20.

[14] Tsay RS, Peña D, Pankratz AE. 2000. Outliers in multivariate time series. *Biometrika* **87, 4**: 789 - 804.

[15] Trabelsi A, Hillmer SC. 1989. A benchmarking approach to forecast combination. *Journal of Business and Economic Statistics* **7**: 353 - 362.

[16] van der Knopp HS. 1987. Conditional forecasting with a multivariate time series model. *Economics Letters* **22**: 233 - 236.

CHAPTER 4

RESTRICTED VAR FORECASTS OF ECONOMIC TIME SERIES WITH CONTEMPORANEOUS CONSTRAINTS

4.1 Introduction

When forecasting economic variables, it frequently happens that some components of a multiple time series must satisfy a contemporaneous binding constraint. Such constraints are intrinsic to macroeconomic and financial data structures underlying the balance of payments, the monetary aggregates, etc. For instance, in the balance of payments account, deficit equals income minus expenditure of the foreign sector. So, the contemporaneous constraint on the bivariate time series is the deficit, which may be considered as a univariate time series by itself. In fact, when planning or implementing economic policies the government usually imposes targets on this aggregated variable.

This paper presents a methodology, within the Vector Auto-Regressive (VAR) framework, for forecasting multivariate time series that satisfy a contemporaneous binding constraint for which there exists a future target. We consider two ways of obtaining the restricted forecasts of the system: (1) the target is introduced directly into the forecasts as a linear combination of the future values of the time series vector; (2) the restricted forecast of the aggregated variable (which becomes the contemporaneous binding constraint of the time series vector) is first obtained and then imposed to get the restricted forecasts of the system. It is well known (see for instance, Pankratz 1989) that restricted forecasts of binding constrained systems are at least as precise as the unrestricted forecasts, besides being internally consistent. It is proved here that the second way of obtaining restricted forecasts, as described above, produces forecasts at least as precise as with the first way. The proposed procedures are illustrated with an example based on the balance of payments account for the Mexican economy. Here, the income and expenditure system of variables is constrained to satisfy a government target for the deficit.

Obtaining restricted forecasts of contemporaneously constrained time series has immediate implications when forecasting cointegrated systems, since a cointegration relationship can be seen as an unbinding contemporaneous constraint. In fact, the restriction is imposed by letting the error correction term be equal to zero at the end of the forecast horizon, so that the dynamical system is assumed to be in equilibrium from that point onwards. By imposing the restriction this way we implicitly define long-run in terms of the time needed for the error correction term to disappear. A Monte Carlo simulation of an artificially generated Vector Error Correction (VEC) model with one unit root was carried out to study the behavior of unrestricted and restricted forecasts. In the forecast comparison, four forecast precision measures were used to judge the results empirically. These results indicate that the restricted forecasts are indeed better than the unrestricted ones.

The restricted forecasting methodology for a multivariate time series has appeared in such works as those of Pankratz (1989) and van der Knoop (1987). In these, the extra-model information was introduced in the form of linear restrictions. Guerrero and Peña (2003) provided general results for the problem of combining data from two different sources of information in order to improve the efficiency of predictors. Some time series problems such as forecast updating when new information is available, forecast combination, interpolation and missing value estimation, among others, can be treated with their proposal. Pandher (2002) attacked the problem of modeling and forecasting a contemporaneously constrained system of time series within the state-space framework. On the other hand, several criteria for comparing forecasts have appeared in the literature. Clements and Hendry (1993) criticized the use of Mean Square Error (MSE) because it is not invariant to non-singular, scale-preserving linear transformations. Lin and Tsay (1996) used the square root of the trace of the covariance matrix of out-of-sample forecasts errors as the main criterion to study the forecasting performance of cointegrated variables with different number of unit roots. Christoffersen and Diebold (1998) dealt with forecasting cointegrated variables and showed that nothing is lost by ignoring cointegration when forecasts are evaluated using the trace of the MSE, since such a measure fails to value the long-run forecasts. Thus, they suggested to use two MSE measures of forecasting performance. The first one is the trace MSE of the cointegrating combinations of the forecast errors. While the second one corresponds to a triangular representation of the cointegrated system that incorporates both the standard MSE and that of the aforementioned cointegrating combinations.

In this paper we proceed as follows. In section 4.2 we establish the notation and standard results of VAR models and the restricted forecasting methodology. In section 4.3, we derive a procedure to get restricted forecasts of a contemporaneously constrained system. In section 4.4, we illustrate the methodology with an empirical application to the balance of payment accounts for the Mexican economy. In section 4.5, we present a Monte Carlo study for an artificially generated VEC model. The results of this study allows us to validate our proposal by way of precision measures usually employed to compare forecasts. We conclude in section 4.6.

4.2 Preliminaries

In this section, we establish the notation and recall some standard results pertaining to the VAR model and to the corresponding restricted forecasting methodology.

4.2.1 Models

Let $\mathbf{y}_t = (y_{1t} \cdots y_{kt})'$ be a $k \times 1$ vector of variables observed at time t, for $t = 1, \ldots, T$ and let us assume that $\mathbf{y}_t$ follows a finite pth-order Gaussian VAR model

$$\mathbf{y}_t = \Lambda \boldsymbol{\delta}_t + \Pi_1 \mathbf{y}_{t-1} + \cdots + \Pi_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \tag{4.1}$$

where Π_i is a $k \times k$ matrix of parameters, for $i = 1, \ldots, p$, $\boldsymbol{\delta}_t = (\delta_{1t} \cdots \delta_{nt})'$ is an $n \times 1$ vector that may include both deterministic variables to account for seasonality, as well as intervention effects and exogenous variables with respect to $\mathbf{y}_t$. $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t} \cdots \varepsilon_{kt})'$ is a $k \times 1$ random error vector independent and identically distributed $N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ with $\boldsymbol{\Sigma}_\varepsilon$ a covariance matrix whose ijth element is $\sigma_{ij} = \text{cov}(\varepsilon_{it}, \varepsilon_{jt})$, for $i, j = 1, 2, \ldots, k$. Thus the $\boldsymbol{\varepsilon}$'s are serially uncorrelated but may be contemporaneously correlated. The effects of $\boldsymbol{\delta}_t$ on $\mathbf{y}_t$ are captured by the $k \times n$ parameter matrix Λ.

Subtracting $\mathbf{y}_{t-1}$ from both sides of (4.1) and rearranging terms we have

$$\Pi^* (B) \Delta \mathbf{y}_t = \Lambda \boldsymbol{\delta}_t - \Pi (1) B \mathbf{y}_t + \boldsymbol{\varepsilon}_t, \tag{4.2}$$

where Δ is the first difference operator, $\Pi(1) = I - \Pi_1 - \cdots - \Pi_p$ and $\Pi^* (B)$ is a matrix polynomial of order $p - 1$ such that $\Pi (B) = \Pi (1) B + \Pi^* (B) \Delta$. We assume the determinantal polynomial $\det(\Pi (x))$ has all its zeros on or outside the unit circle and the rank of $\Pi (1)$ is r, implying that there are $d = k - r$ unit roots in the system. Moreover, we assume the unit roots have multiplicity one, meaning that $\mathbf{y}_t$ is at most integrated of order 1, which is written as $I(1)$. When $r > 0$ the variables are cointegrated, in the sense that there exists a linear combination $\boldsymbol{\beta}\mathbf{y}_t$, with $\boldsymbol{\beta} = (\beta_1 \cdots \beta_k) \neq \mathbf{0}$, which is $I(0)$. In that case the matrix $-\Pi(1) = \boldsymbol{\alpha}\boldsymbol{\beta}$ where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $k \times r$ and $r \times k$ full rank matrices respectively. Equation (4.2) then becomes

$$\Pi^*(B)\,\Delta\mathbf{y}_t = \Lambda\boldsymbol{\delta}_t + \boldsymbol{\alpha}z_{t-1} + \boldsymbol{\varepsilon}_t, \tag{4.3}$$

where $z_t = \boldsymbol{\beta}\mathbf{y}_t$ is stationary. The rows of $\boldsymbol{\beta}$ are then referred as cointegration vectors of the system and equation (4.3) is known as the Vector Error Correction (VEC) representation of (4.1).

4.2.2 Forecasts

Let us assume that $\mathbf{Y} = (\mathbf{y}_1' \cdots \mathbf{y}_T')'$ is a $kT \times 1$ vector containing all the past information of the multiple time series and let $\mathbf{Y}_F = \left(\mathbf{y}_{T+1}' \cdots \mathbf{y}_{T+H}'\right)'$ denote a $kH \times 1$ vector containing the $H \geq 1$ values to be forecasted for each series.

The optimal (in MSE sense) linear forecast of $\mathbf{y}_{T+h}$, for $h = 1, \ldots, H$, is its conditional expectation

$$\mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \Lambda\boldsymbol{\delta}_{T+h} + \Pi_1\mathrm{E}(\mathbf{y}_{T+h-1}|\mathbf{Y}) + \cdots + \Pi_p\mathrm{E}(\mathbf{y}_{T+h-p}|\mathbf{Y}) \tag{4.4}$$

where $\mathrm{E}(\mathbf{y}_{T+h-i}|\mathbf{Y}) = \mathbf{y}_{T+h-i}$ for $i \geq h$. Such a forecast produces the h-step-ahead forecast error vector

$$e_T \equiv \mathbf{y}_{T+h} - \mathrm{E}(\mathbf{y}_{T+h}|\mathbf{Y}) = \sum_{j=0}^{h-1}\Psi_j\boldsymbol{\varepsilon}_{T+h-j}, \text{ for } h = 1, \ldots, H \tag{4.5}$$

where $\Psi_j = \sum_{k=1}^{j}\Psi_{j-k}\Pi_k$ with $\Psi_0 = I$ and $\Pi_k = 0$ for $k > p$, see Lütkepohl (1991, Section 11.3). It should be stressed that, by virtue of the equivalence between (4.1) and (4.2), the forecasts provided by (4.4) make use of all the information in the VEC model. Stacking the forecast errors (4.5) we write

$$\mathbf{e}_F \equiv \mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \boldsymbol{\Psi}\boldsymbol{\epsilon}_F$$

where $\boldsymbol{\epsilon}_F = \left(\boldsymbol{\varepsilon}'_{T+1}\cdots\boldsymbol{\varepsilon}'_{T+H}\right)' \sim N\left(\mathbf{0}, I_H \otimes \boldsymbol{\Sigma}_\varepsilon\right)$ is a $kH \times 1$ random vector, with $\otimes$ the Kronecker product and $\boldsymbol{\Psi}$ is a $kH \times kH$ lower triangular matrix with Ψ_0 in its main diagonal, Ψ_1 in the first subdiagonal, Ψ_2 in the second subdiagonal and so on.

The MSE of $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})$ is given by

$$\begin{aligned}
\Sigma_{\mathbf{Y}} &\equiv \mathrm{E}[(\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}))(\mathbf{Y}_F - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}))'|\mathbf{Y}], \\
&= \boldsymbol{\Psi}\left(I_H \otimes \boldsymbol{\Sigma}_\varepsilon\right)\boldsymbol{\Psi}'.
\end{aligned}$$

When forecasting, we sometimes have additional information about the future that cannot be included directly into the model. For instance, government economic targets provide additional information, or extra-model information as Pankratz (1989) called it, that cannot be included for estimation purposes, but should be taken into account when making forecasts to reflect the idea that some policy intervention is expected to occur. We want to introduce this additional information into the forecast of $\mathbf{y}_t$ and assume that it is in the form of the following stochastic linear restriction

$$\mathbf{R} = \mathbf{C}\mathbf{Y}_F + \mathbf{u}. \tag{4.6}$$

Here, $\mathbf{C}$ is an $M \times kH$ constant matrix containing the coefficients that define the linear combinations of $\mathbf{Y}_F$,

$$\mathbf{C} = \left(\begin{array}{ccccccc} c_{1,1} & \cdots & c_{1,k} & \cdots & c_{1,k(H-1)+1} & \cdots & c_{1,kH} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ c_{M,1} & \cdots & c_{M,k} & \cdots & c_{M,k(H-1)+1} & \cdots & c_{M,kH} \end{array} \right)$$

$$\underbrace{\phantom{c_{M,1} \cdots c_{M,k}}}_{\mathbf{y}_{T+1}} \qquad \underbrace{\phantom{c_{M,k(H-1)+1} \cdots c_{M,kH}}}_{\mathbf{y}_{T+H}}$$

where the rows are linearly independent, so that the rank of $\mathbf{C}$ is M. Besides, $\mathbf{R} = (r_1 \cdots r_M)'$ is a $M \times 1$ vector of values that the linear combinations are expected to take on and $\mathbf{u}$ is an $M \times 1$ random vector such that $\mathbf{u} = (u_1 \cdots u_M)' \sim N(\mathbf{0}, \Sigma_{\mathbf{u}})$, with the ijth element of $\Sigma_{\mathbf{u}}$ given by $\sigma_{ij,u}$ =cov(u_i, u_j), for $i, j = 1, \ldots, M$.

It can be shown that the optimal restricted forecast, in MSE sense, is given by

$$\widehat{\mathbf{Y}}_F = \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) + \mathbf{A}\left[\mathbf{R} - \mathbf{C}\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})\right], \tag{4.7}$$

where $\mathbf{A} = \Sigma_{\mathbf{Y}}\mathbf{C}'\left(\mathbf{C}\Sigma_{\mathbf{Y}}\mathbf{C}' + \Sigma_{\mathbf{u}}\right)^{-1}$ (see for instance, Pankratz 1989, for a proof) and its MSE is given by

$$\mathrm{MSE}\left(\widehat{\mathbf{Y}}_F\right) = (I - \mathbf{AC})\,\Sigma_{\mathbf{Y}}. \tag{4.8}$$

For A and B two symmetric matrices we write $A \geq B$ if $A - B$ is positive semidefinite and $A > B$ if $A - B$ is positive definite. Hence, $\Sigma_{\mathbf{Y}} \geq$MSE$\left(\widehat{\mathbf{Y}}_F\right)$ since $\mathbf{AC}\Sigma_{\mathbf{Y}}$ is a positive semidefinite matrix, and we interpret this result as saying that $\widehat{\mathbf{Y}}_F$ is at least as precise as $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})$.

4.3 Restricted forecasts for contemporaneously constrained VAR models

When studying economic variables, it some times happens that certain components of a time series vector must satisfy a contemporaneous binding constraint. Such a constraint might come from macroeconomic or financial identities. Here we assume this constraint is a linear combination of the elements of $\mathbf{y}_t$, that is $z_t \equiv \mathbf{c}\mathbf{y}_t$. When this happens, we shall call z_t the *contemporaneous constraint* and $\mathbf{y}_t$ the *contemporaneously constrained vector*. Let $\mathbf{Z} = (z_1 \cdots z_T)'$ be a $T \times 1$ vector containing all the past information of z_t and let $\mathbf{Z}_F = (z_{T+1} \cdots z_{T+H})'$ denote an $H \times 1$ vector that contains $H \geq 1$ future values of $\{z_t\}$.

When planning or implementing economic policies it is common to impose targets on the aggregated variable z_t. For example, in the Mexican foreign sector the economic targets are imposed on the deficit. So, the problem of incorporating additional information into the forecasts of a contemporaneously constrained system arises. To solve this problem, we assume the additional information is binding and related to the future values of z_t, so that equation (4.6) becomes

$$r = \mathbf{C}_0 \mathbf{Z}_F, \tag{4.9}$$

where $\mathbf{C}_0$ contains the coefficients of the linear combination of $\mathbf{Z}_F$.

In this setup, the restricted forecasts of $\mathbf{y}_t$ can be obtained (1) by introducing the additional information directly into the forecasts. We shall call this the *one-stage restricted forecast*. Or (2) by forecasting the univariate time series z_t

restricted by the additional information as a first step, and then introducing this *univariate restricted forecast* as a restriction on the system, in a second step. We will refer to the second approach as the *two-stage restricted forecast.* The arrays involved with the univariate restricted forecasts will be denoted with the subindex 0. These forecasts are obtained as follows.

Univariate restricted forecast. When $k = 1$, equations (4.7) and (4.8) provide the restricted forecasts for an AR(p) process. Fortunately, the formulas stay the same for ARIMA (Auto-Regressive Integrated and Moving Average) processes, as it was proved in Guerrero (1989). We will take advantage of this fact and allow for the use of ARIMA processes, since these are dynamically richer than the AR processes. So, the restricted forecast of the vector $\mathbf{Z}_F$ and its MSE will be denoted by

$$\widehat{\mathbf{Z}}_F = \mathrm{E}(\mathbf{Z}_F|\mathbf{Y}) + \mathbf{A}_0 \left[r - \mathbf{C}_0\mathrm{E}(\mathbf{Z}_F|\mathbf{Y})\right] \tag{4.10}$$

$$\mathrm{MSE}\left(\widehat{\mathbf{Z}}_F\right) = \left(I - \mathbf{A}_0\mathbf{C}_0\right)\Sigma_{\mathbf{Z}}, \tag{4.11}$$

where $\Sigma_{\mathbf{Z}}$ =MSE[E($\mathbf{Z}_F|\mathbf{Y}$)] and $\mathbf{A}_0 = \Sigma_{\mathbf{Z}}\mathbf{C}_0' \left(\mathbf{C}_0\Sigma_{\mathbf{Z}}\mathbf{C}_0'\right)^{-1}$. Let us notice in particular that $\Sigma_{\mathbf{Z}} \geq \mathrm{MSE}\left(\widehat{\mathbf{Z}}_F\right)$.

(1) **One-stage restricted forecast**. The additional information can be written in terms of $\mathbf{Y}_F$ by noticing that $r = \mathbf{C}_0\mathbf{Z}_F = \mathbf{C}_0\left(\mathbf{c}\mathbf{y}_{T+1}\cdots\mathbf{c}\mathbf{y}_{T+H}\right)' = \mathbf{C}_1\mathbf{Y}_F$, where $\mathbf{C}_1 = \mathbf{C}_0 \otimes \mathbf{c}$ is a $1 \times kH$ matrix. Hence, the multivariate restricted forecast of $\mathbf{Y}_F$ and its MSE will be given by

$$\widehat{\mathbf{Y}}_{F,1} = \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) + \mathbf{A}_1 \left[r - \mathbf{C}_1\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})\right] \tag{4.12}$$

$$\text{MSE}\left(\widehat{\mathbf{Y}}_{F,1}\right) = \left(I - \mathbf{A}_1\mathbf{C}_1\right)\Sigma_{\mathbf{Y}}, \tag{4.13}$$

where $\mathbf{A}_1 = \Sigma_{\mathbf{Y}}\mathbf{C}_1'\left(\mathbf{C}_1\Sigma_{\mathbf{Y}}\mathbf{C}_1'\right)^{-1}$ and $\Sigma_{\mathbf{Y}}$ is as defined previously.

(2) **Two-stage restricted forecast**. Another (more efficient) way to obtain the multivariate restricted forecast consists of using the fact that $\widehat{\mathbf{Z}}_F$ already satisfies the additional information ($\mathbf{C}_0\widehat{\mathbf{Z}}_F = r$). So, if the forecast of $\mathbf{Y}_F$ is now restricted by the univariate restricted forecast $\widehat{\mathbf{Z}}_F$ with forecast error covariance matrix given by MSE$\left(\widehat{\mathbf{Z}}_F\right)$, then the stochastic linear restriction (4.6) becomes

$$\widehat{\mathbf{Z}}_F = \mathbf{C}_2\mathbf{Y}_F + \mathbf{v}; \quad \mathbf{v} \sim N\left(\mathbf{0}, \Sigma_{\mathbf{v}}\right)$$

where $\mathbf{C}_2 = I_H \otimes \mathbf{c}$ is an $H \times kH$ matrix representing the contemporaneous constraint $\widehat{\mathbf{Z}}_F - \mathbf{v} = \mathbf{C}_2\mathbf{Y}_F$ (which can be interpreted as constraining the elements of $\mathbf{Y}_F$ to satisfy $\widehat{\mathbf{Z}}_F$ plus a forecast error) and $\Sigma_{\mathbf{v}}$ =MSE$\left(\widehat{\mathbf{Z}}_F\right)$. The restricted forecast and its MSE for this case are

$$\widehat{\mathbf{Y}}_{F,2} = \text{E}(\mathbf{Y}_F|\mathbf{Y}) + \mathbf{A}_2\left[\widehat{\mathbf{Z}}_F - \mathbf{C}_2\text{E}(\mathbf{Y}_F|\mathbf{Y})\right] \tag{4.14}$$

$$\text{MSE}\left(\widehat{\mathbf{Y}}_{F,2}\right) = \left(I - \mathbf{A}_2\mathbf{C}_2\right)\Sigma_{\mathbf{Y}} \tag{4.15}$$

where $\mathbf{A}_2 = \Sigma_{\mathbf{Y}}\mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' + \Sigma_{\mathbf{v}}\right)^{-1}$.

The following result simply states that the two-stage restricted forecast is at least as precise as the one-stage restricted forecast.

Proposition 2 *If the $k \times 1$ vector time series $\{\mathbf{y}_t\}$ given by (4.1) is contemporaneously constrained for $t = 1, \ldots, T$ and the additional information is given by (4.9), then MSE$\left(\widehat{\mathbf{Y}}_{F,1}\right) \geq$MSE$\left(\widehat{\mathbf{Y}}_{F,2}\right)$.*

Proof. First we see that

$$\begin{aligned}
\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,1}\right) - \mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,2}\right) &= \left(I - \mathbf{A}_1\mathbf{C}_1\right)\Sigma_{\mathbf{Y}} - \left(I - \mathbf{A}_2\mathbf{C}_2\right)\Sigma_{\mathbf{Y}} \\
&= \left(\mathbf{A}_2\mathbf{C}_2 - \mathbf{A}_1\mathbf{C}_1\right)\Sigma_{\mathbf{Y}} \\
&= \Sigma_{\mathbf{Y}}\left[\mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' + \Sigma_{\mathbf{v}}\right)^{-1}\mathbf{C}_2 - \mathbf{C}_1'\left(\mathbf{C}_1\Sigma_{\mathbf{Y}}\mathbf{C}_1'\right)^{-1}\mathbf{C}_1\right]\Sigma_{\mathbf{Y}}.
\end{aligned}$$

Thus, the problem reduces to prove that

$$\mathbf{\Omega} \equiv \mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' + \Sigma_{\mathbf{v}}\right)^{-1}\mathbf{C}_2 - \mathbf{C}_1'\left(\mathbf{C}_1\Sigma_{\mathbf{Y}}\mathbf{C}_1'\right)^{-1}\mathbf{C}_1 \tag{4.16}$$

is a positive semidefinite matrix.

By applying the Matrix Inversion Lemma (see Harvey 1993, page 104) and noticing that $\mathbf{C}_1 = \mathbf{C}_0\mathbf{C}_2$, we have

$$\begin{aligned}
\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' + \Sigma_{\mathbf{v}}\right)^{-1} &= \left[2\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' - \mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\mathbf{C}_0'\left(\mathbf{C}_0\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\mathbf{C}_0'\right)^{-1}\mathbf{C}_0\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right]^{-1} \\
&= \frac{1}{2}\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1} + \frac{1}{2}\mathbf{C}_0'\left(\mathbf{C}_0\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\mathbf{C}_0'\right)^{-1}\mathbf{C}_0
\end{aligned}$$

and from here we get

$$\mathbf{\Omega} = \frac{1}{2}\left[\mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\mathbf{C}_2 - \mathbf{C}_1'\left(\mathbf{C}_1\Sigma_{\mathbf{Y}}\mathbf{C}_1'\right)^{-1}\mathbf{C}_1\right]. \tag{4.17}$$

Thus, from equations (4.16) and (4.17), and applying again the Matrix Inversion Lemma, we obtain

$$\begin{aligned}
\mathbf{\Omega} &= \mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\mathbf{C}_2 - \mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2' + \Sigma_{\mathbf{v}}\right)^{-1}\mathbf{C}_2 \\
&= \mathbf{C}_2'\left[\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\left(\Sigma_{\mathbf{v}}^{-1} + \left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\right)^{-1}\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\right]\mathbf{C}_2 \quad (4.18)
\end{aligned}$$

which is clearly a positive semidefinite matrix. □

The following proposition states a condition for equality of the one-stage and two-stage restricted forecasts.

Proposition 3 *If the $k \times 1$ vector time series $\{\mathbf{y}_t\}$ given by (4.1) is contemporaneously constrained for $t = 1, \ldots, T$, with additional information given by (4.9) and $\Sigma_{\mathbf{v}} = \mathbf{0}$, then $\widehat{\mathbf{Y}}_{F,1} = \widehat{\mathbf{Y}}_{F,2}$ and $MSE\left(\widehat{\mathbf{Y}}_{F,1}\right) = MSE\left(\widehat{\mathbf{Y}}_{F,2}\right)$.*

Proof. Equation (4.10) can be written as

$$\widehat{\mathbf{Z}}_F - \mathbf{C}_2 \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \mathbf{A}_0 \left[r - \mathbf{C}_1 \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) \right]$$

since $\mathbf{Z}_F = \mathbf{C}_2 \mathbf{Y}_F$ and $\mathbf{C}_1 = \mathbf{C}_0 \mathbf{C}_2$. Now, from this equation and (4.14) we have

$$\widehat{\mathbf{Y}}_{F,2} - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \mathbf{A}_2 \mathbf{A}_0 \left[r - \mathbf{C}_1 \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) \right].$$

Then, since $\mathbf{A}_0 = \mathbf{C}_2 \Sigma_{\mathbf{Y}} \mathbf{C}_2' \mathbf{C}_0' \left(\mathbf{C}_1 \Sigma_{\mathbf{Y}} \mathbf{C}_1' \right)^{-1}$ and $\mathbf{A}_2 = \Sigma_{\mathbf{Y}} \mathbf{C}_2' \left(\mathbf{C}_2 \Sigma_{\mathbf{Y}} \mathbf{C}_2' + \Sigma_{\mathbf{v}} \right)^{-1}$ we get

$$\widehat{\mathbf{Y}}_{F,2} - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \Sigma_{\mathbf{Y}} \mathbf{C}_2' \left(\mathbf{C}_2 \Sigma_{\mathbf{Y}} \mathbf{C}_2' + \Sigma_{\mathbf{v}} \right)^{-1} \mathbf{C}_2 \Sigma_{\mathbf{Y}} \mathbf{C}_2' \mathbf{C}_0' \left(\mathbf{C}_1 \Sigma_{\mathbf{Y}} \mathbf{C}_1' \right)^{-1} \left[r - \mathbf{C}_1 \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) \right]. \tag{4.19}$$

On the other hand, from (4.12) we know that

$$\widehat{\mathbf{Y}}_{F,1} - \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) = \Sigma_{\mathbf{Y}} \mathbf{C}_1' \left(\mathbf{C}_1 \Sigma_{\mathbf{Y}} \mathbf{C}_1' \right)^{-1} \left[r - \mathbf{C}_1 \mathrm{E}(\mathbf{Y}_F|\mathbf{Y}) \right].$$

Therefore, when $\Sigma_{\mathbf{v}} = \mathbf{0}$, we have $\widehat{\mathbf{Y}}_{F,1} = \widehat{\mathbf{Y}}_{F,2}$, and $\mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,1}\right) = \mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,2}\right)$ as indicated by expression (4.18). □

Some remarks are now in order:

(i) If conditions of proposition 3 hold, we know $\widehat{\mathbf{Y}}_{F,1} = \widehat{\mathbf{Y}}_{F,2}$, but it is preferable (by simplicity) to calculate the one-stage rather than the two-stage restricted forecast.

(ii) By applying the Matrix Inversion Lemma to expression (4.19) we get

$$\begin{aligned}\widehat{\mathbf{Y}}_{F,2} &= \widehat{\mathbf{Y}}_{F,1} - \Sigma_{\mathbf{Y}}\mathbf{C}_2'\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1}\left[\left(\mathbf{C}_2\Sigma_{\mathbf{Y}}\mathbf{C}_2'\right)^{-1} + \Sigma_{\mathbf{v}}^{-1}\right]^{-1}\mathbf{C}_0'\left(\mathbf{C}_1\Sigma_{\mathbf{Y}}\mathbf{C}_1'\right)^{-1} \times \\ & \quad [r - \mathbf{C}_1\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})]\end{aligned}$$

which says that $\widehat{\mathbf{Y}}_{F,2}$ can be obtained by first calculating $\widehat{\mathbf{Y}}_{F,1}$ and then adding a term that depends on $\Sigma_{\mathbf{v}}$.

(iii) The following forecast efficiency inequality holds true

$$\Sigma_{\mathbf{Y}} \geq \mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,1}\right) \geq \mathrm{MSE}\left(\widehat{\mathbf{Y}}_{F,2}\right). \tag{4.20}$$

4.4 An empirical illustration with the balance of payments account

The variables employed in this exercise are *Current Account Deficit* **(DEF)**, defined as income minus expenditure of the foreign sector, *Income of the Foreign Sector* **(INCM)** and *Expenditure of the Foreign Sector* **(EXPN)**. When the Mexican Government published the economic targets for 2005 the data available ran from 1980:1 to 2004:2. Thus, the data set consists of 98 quarterly observations on each series expressed in millions of dollars (the data come from Bank of Mexico, http://www.banxico.org.mx/).

4.4.1 Order of integration

The order of integration of the series was decided on the basis of Augmented Dickey-Fuller (ADF) tests. The regression model included a constant term, a deterministic trend for INCM and EXPN and centered dummies for all the variables. We included three dummy variables, two for the first and second quarters of 1995 to account for the Mexican crisis, and one more for the first quarter of 2001. The general equation employed was

$$\Delta x_t = a + b_0 t + \sum_{i=1}^{n} b_i \delta_{it} + c_0 x_{t-1} + \sum_{j=1}^{p} c_j \Delta x_{t-j} + \varepsilon_t.$$

Table 4.1 shows the results of the ADF tests with and without the dummy variables. The τ statistic allows us to test H_0:$c_0 = 0$. The symbol (*) indicates rejection of the null hypothesis at the 5% significance level. Critical values do not consider the presence of dummy variables. The order, p, of the autoregression was selected to guarantee no residual autocorrelation. All variables turned out to be $I(1)$ and the order of integration does not depend on the inclusion of dummy variables.

4.4.2 Model estimation

To illustrate the proposed methodology we estimated a model for the current account deficit and another one for the income-expenditure system. The income-expenditure system required also three dummy variables, two for the first and second quarters of 1995 and another one for 2001:1.

Table 4.1: Unit root ADF results

Variable	$H_0 : I(1)$				$H_0 : I(2)$			
	Without dummies		With dummies		Without dummies		With dummies	
	p	τ	p	τ	p	τ	p	τ
DEF	3	-2.19	4	-1.59	3	$-4.18*$	3	$-4.46*$
INCM	0	-2.21	0	-1.93	0	$-10.28*$	0	$-10.35*$
EXPN	2	-3.20	2	-2.95	4	$-5.03*$	2	$-4.02*$

4.4.2.1 Deficit variable

A seasonal moving average model (with seasonality of order 4) was found adequate for the current account deficit with regular and seasonal differencing, including two dummy variables for the first and second quarters of 1995

$$\Delta\Delta_4 z_t = \underset{(4.1)}{3324.5} I_{95:1} + \underset{(3.2)}{2606.8} I_{95:2} + (1 - \underset{(-9.7)}{0.742} B^4)\varepsilon_t; \qquad \widehat{\sigma}_\varepsilon = 1049.04.$$

The numbers in parentheses are t-statistics. Figure 4.1 presents the observed and estimated series, together with the residuals.

4.4.2.2 Income-expenditure system

The income and expenditure variables turned out to be not cointegrated (see the Johansen tests results in Table 4.2). Thus, a VAR model for $\Delta \mathbf{y}_t = (\Delta\mathrm{INCM}_t, \Delta\mathrm{EXPN}_t)'$ will provide a reasonable representation for the system. The AIC criterion led us to use $p = 4$ as the order of the autoregression. The

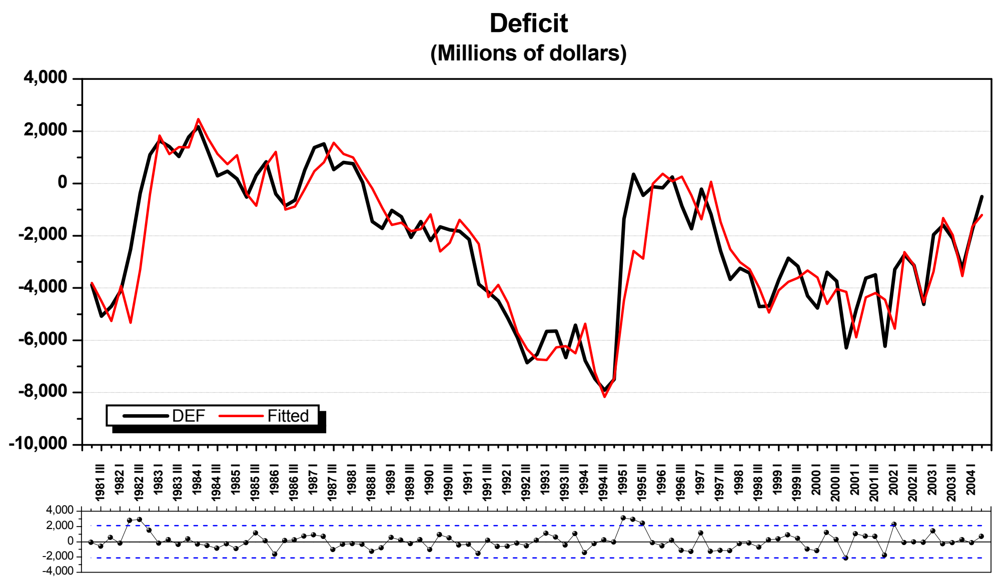


Figure 4.1: Observed and fitted series in levels at the top, and corresponding residuals with $\pm 2\widehat{\sigma}$ horizontal lines, at the bottom

VAR model also included a constant term, centered dummy variables to account for seasonal effects, two dummy variables to account for the 1995 economic crisis and one more dummy for 2001:1. Hence, the deterministic vector became $\boldsymbol{\delta}_t = (\text{const}, S_1, S_2, S_3, I_{95:1}, I_{95:2}, I_{01:1})'$ and the estimation of equation (4.1) produced the following matrices

$$\widehat{\Lambda} = \begin{pmatrix} \underset{(2.19)}{294.8} & \underset{(-4.18)}{-1741.4} & \underset{(-0.25)}{-93.7} & \underset{(-1.21)}{-486.1} & \underset{(1.95)}{2133.6} & \underset{(-2.33)}{-3012.7} & \underset{(-3.36)}{-3890.8} \\ \underset{(1.35)}{226.0} & \underset{(-2.69)}{-1396.4} & \underset{(0.19)}{88.2} & \underset{(0.52)}{258.3} & \underset{(-3.02)}{-4108.2} & \underset{(-2.87)}{-4615.7} & \underset{(-4.20)}{-6052.6} \end{pmatrix},$$

$$\widehat{\Pi}_1 = \begin{pmatrix} \underset{(3.39)}{0.544} & \underset{(-3.37)}{-0.424} \\ \underset{(2.54)}{0.507} & \underset{(-1.82)}{-0.285} \end{pmatrix}, \quad \widehat{\Pi}_2 = \begin{pmatrix} \underset{(0.45)}{0.064} & \underset{(2.59)}{0.279} \\ \underset{(2.13)}{0.373} & \underset{(1.02)}{0.137} \end{pmatrix},$$

$$\widehat{\Pi}_3 = \begin{pmatrix} \underset{(0.13)}{0.019} & \underset{(0.23)}{0.024} \\ \underset{(-0.59)}{-0.105} & \underset{(0.18)}{0.024} \end{pmatrix}, \quad \widehat{\Pi}_4 = \begin{pmatrix} \underset{(1.94)}{0.266} & \underset{(-2.04)}{-0.223} \\ \underset{(1.37)}{0.233} & \underset{(-0.02)}{-0.002} \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}_\varepsilon = \begin{pmatrix} 948.8^2 & 815747.8 \\ 815747.8 & 1181.1^2 \end{pmatrix}.$$

with $R^2 = 0.58$ and $R^2 = 0.61$ for income and expenditure, respectively. Figure 4.2 shows the observed and estimated series of the income-expenditure system.

Table 4.2: Johansen cointegration analysis

Null	Trace Statistic	Crit 90%	Crit 95%	Eigen Statistic	Crit 90%	Crit 95%
$r \leq 0$	10.53	13.43	15.49	8.59	12.30	14.26
$r \leq 1$	1.94	2.71	3.84	1.94	2.71	3.84

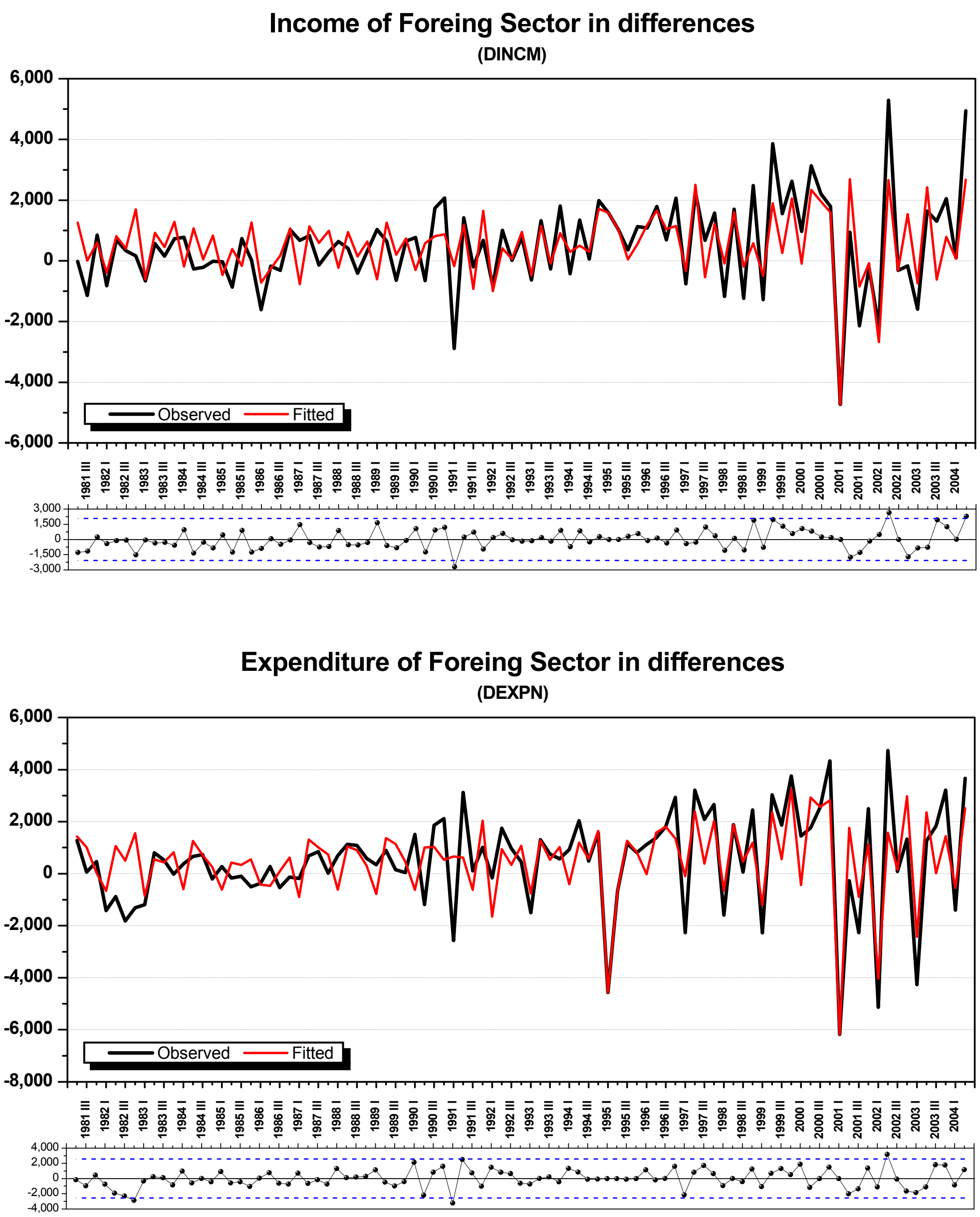


Figure 4.2: Observed and fitted series in first differences, and residuals with horizontal lines corresponding to $\pm 2\widehat{\sigma}$

4.4.3 Restricted forecasts

At the end of year 2004, the Mexican Government published the economic targets for 2005 (see SHCP, 2004). It was foreseen that the current account deficit will move from -8,887.8 (-1.4% of GDP in 2004) to -14,237.4 (-2.1% of GDP in 2005).

4.4.3.1 Univariate restricted forecast of deficit

The above economic target can be written as

$$\mathrm{DEF}_{2005:1} + \mathrm{DEF}_{2005:2} + \mathrm{DEF}_{2005:3} + \mathrm{DEF}_{2005:4} = -14,237.4$$

Since the current public administration will stay in power up to the end of year 2006, it was natural to choose the forecast horizon as $H = 10$. Thus the linear stochastic restriction (4.9) for $H = 10$ periods ahead, gets specified by means of

$$r = (-14,237.4) \text{ and } \mathbf{C}_0 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence the univariate restricted forecast and its MSE are given by equations (4.10) and (4.11). Figure 4.3 shows the unrestricted and restricted forecasts, with their 90% probability intervals.

4.4.3.2 Income-expenditure system, one-stage restricted forecast

Since the current account deficit is defined as income minus expenditure of the foreign sector, the additional information to get a forecast with $H = 10$ periods

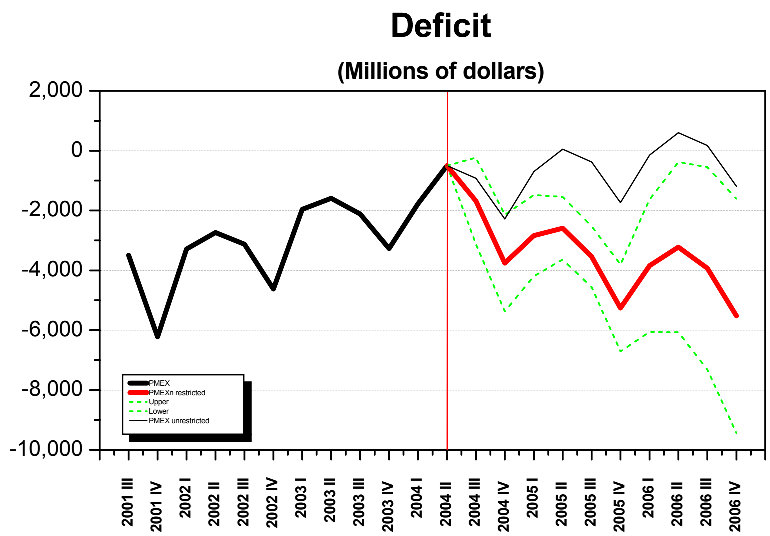


Figure 4.3: Restricted forecasts with 90% probability intervals (origin at 2004:2)

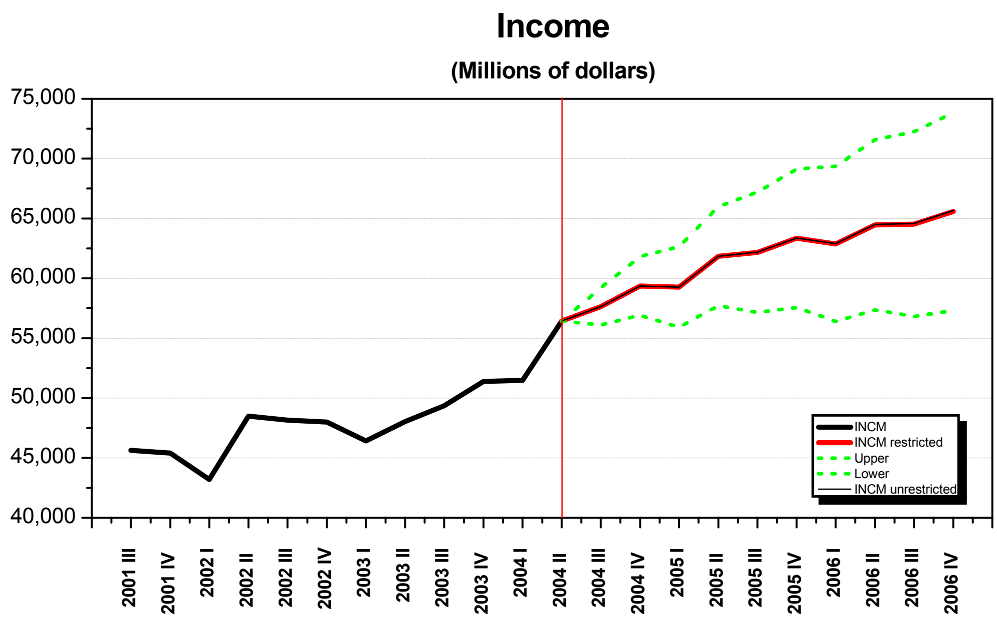
ahead gets determined by

$$r = (-14, 237.4) \text{ and } \mathbf{C}_1 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{c} & \mathbf{c} & \mathbf{c} & \mathbf{c} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{0}$ is the 1×2 zero vector and $\mathbf{c} = \begin{pmatrix} 1 & -1 \end{pmatrix}$. The restricted forecast as well as its MSE are given by equations (4.12) and (4.13). Figure 4.4 shows the unrestricted and one-stage restricted forecast paths for the two variables, with their 90% probability intervals.

4.4.3.3 Income-expenditure system, two-stage restricted forecast

Since the univariate restricted forecasts obtained for deficit take into account the economic target, they can be introduced directly as future contemporaneous



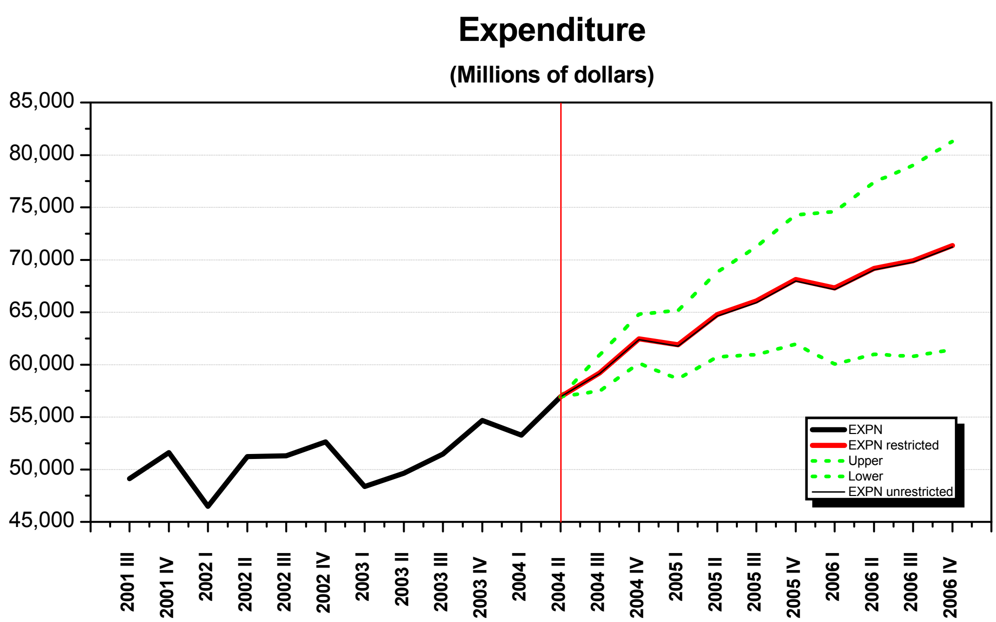


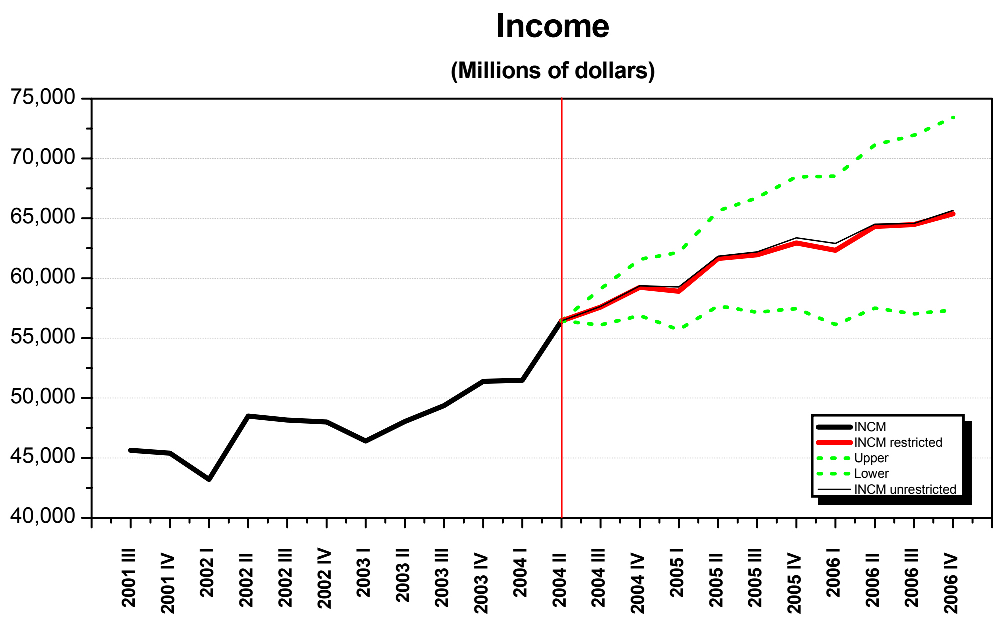
Figure 4.4: One-stage restricted forecasts with 90% probability intervals (origin at 2004:2)

restriction for the system. Thus the additional information to get the forecasts is given by $\widehat{\mathbf{Z}}_F$, $\Sigma_{\mathbf{v}}$ and $\mathbf{C}_2 = I \otimes \mathbf{c}$ where $\mathbf{c}$ is as previously defined.

The two-stage restricted forecast of the system as well as its MSE are given by equations (4.14) and (4.15). Figure 4.5 shows the corresponding paths of the unrestricted and restricted forecasts for the two variables, with their 90% probability intervals. There we corroborate that the two-stage method is more efficient than the one-stage method empirically. For instance, the difference in length of the probability intervals for INCM at the end of the forecast horizon is 433 million dollars, and that for EXPN is 1749 million dollars, both in favor of the two-stage method.

4.5 An extension to cointegrated systems

The restricted forecasting methodology for contemporaneously constrained VAR models of Section 4.3 has immediate implications when forecasting cointegrated systems with one unit root. In that case, a cointegration relationship may be considered as an unbinding contemporaneous constraint. That is, the error correction term $z_t = \boldsymbol{\beta}\mathbf{y}_t$ becomes a contemporaneous constraint on the time series vector $\{\mathbf{y}_t\}$ through the cointegration vector $\boldsymbol{\beta}$. When planning or implementing policies it could be convenient to fix a finite point in the future (say $T+H$) as the moment when the equilibrium will be reached. Hence, the equilibrium restriction on this contemporaneous constrained system can be imposed on the forecast by



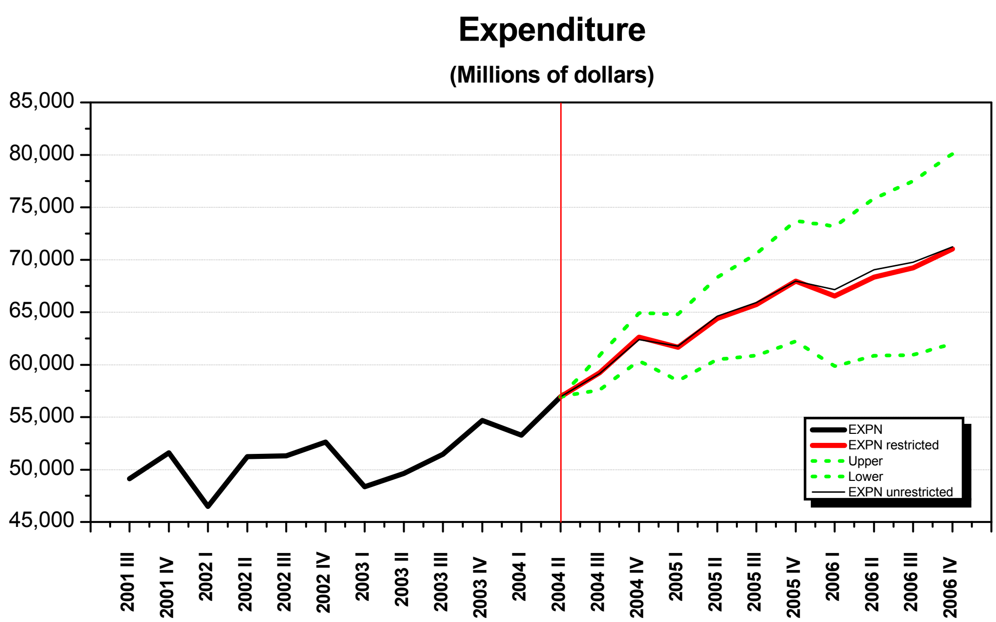


Figure 4.5: Two-stage restricted forecasts with 90% probability intervals (origin at 2004:2)

letting the error correction term be equal to zero at the end of the forecast horizon, so that, $z_{T+H} = 0$.

In this context, the basic aim of this work is to compare, for a one unit root cointegrated system, the forecast efficiency for the unrestricted, one-stage restricted and two-stage restricted forecasts, see expression (4.20). To that end, we obtained N replications of a bivariate process (described below) and estimate it to get the precision of out-of-sample forecasts measured as the square root of the trace of the covariance matrix of the one-step-ahead forecast errors in the forecast horizon (following an approach similar to Lin and Tsay's, 1996),

$$E(H) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\text{trace}\widehat{\Sigma}_i} \tag{4.21}$$

where $\widehat{\Sigma}_i$ is the standard MSE for the ith replication.

However, precision measures that use the standard MSE have been severely criticized. For instance, Clements and Hendry (1993) found that MSE is not invariant to non-singular, scale-preserving linear transformations and Christoffersen and Diebold (1998) pointed out that while the standard MSE measure fails to value the long-run forecasts, the problem with the MSE of the cointegrated combinations is that it values only the long-run forecasts. Due to this fact, they proposed to compute the MSE of the triangular system that, in some way, takes into account the two previous MSE measures.

Thus, we decided to judge the forecasting simulation results of the bivariate system according to four MSE measures. So, $\widehat{\Sigma}_i$ is given by one of the following four criteria:

(1) The *model-based MSE* formulas derived in sections 4.2 and 4.3.

(2) The *standard MSE* defined as

$$\mathrm{MSE} = \mathbf{e}_F \mathbf{e}_F'$$

where $\mathbf{e}_F$ is a $2H \times 1$ vector containing the first H one-step-ahead forecast errors of $\mathbf{Y}_F$.

(3) The *MSE of cointegrating combinations* defined as

$$\mathrm{MSE} = \left[(I \otimes \boldsymbol{\beta})\, \mathbf{e}_F \right] \left[(I \otimes \boldsymbol{\beta})\, \mathbf{e}_F \right]' . \tag{4.22}$$

where $\boldsymbol{\beta}$ is a 1×2 matrix whose row is the cointegration vector of the system.

(4) The *MSE of the triangular system* as given by

$$\mathrm{MSE} = (\Theta \mathbf{u}_F)\,(\Theta \mathbf{u}_F)' ;\ \text{with } \Theta = \begin{bmatrix} \mathbf{0}_{1\times 2} & \boldsymbol{\beta} & & & & \\ -v & v & & & & \\ & \mathbf{0}_{1\times 2} & \boldsymbol{\beta} & & & \\ & -v & v & & & \\ & & & \ddots & & \\ & & & & \mathbf{0}_{1\times 2} & \boldsymbol{\beta} \\ & & & & -v & v \end{bmatrix} \tag{4.23}$$

where $\mathbf{u}_F = (\widehat{\varepsilon}_T, \mathbf{e}_F)'$ is a $2(H+1) \times 1$ vector, $\widehat{\varepsilon}_T$ is the 2×1 vector of residuals at time T, $v = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and Θ is a $2H \times 2(H+1)$ matrix. The other elements of Θ not included in the diagonal shown in (4.23) are all zeros.

To study the performance of the unrestricted and restricted forecasts, a Monte Carlo experiment was done, so that the precision measures of these forecasts were compared with the four MSE measures defined previously. The numerical simulations were generated with the following VEC process

$$\Delta \mathbf{y}_t = \Lambda + \boldsymbol{\alpha}\boldsymbol{\beta}\mathbf{y}_{t-1} + \Pi_1^* \Delta \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{4.24}$$

where $\Lambda = (0 \quad 0)'$, $\boldsymbol{\alpha} = (0.2 \quad 0.1)'$, $\boldsymbol{\beta} = (1 \quad -4)$, $\Pi_1^* = \begin{pmatrix} -0.40 & 0.25 \\ 0.10 & -0.25 \end{pmatrix}$ and $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ with $\boldsymbol{\Sigma}_\varepsilon = \begin{pmatrix} 9 & 0 \\ 0 & 36 \end{pmatrix} \times 10^{-4}$.

The corresponding VAR representation is given by

$$\mathbf{y}_t = \begin{pmatrix} 0.80 & -0.55 \\ 0.20 & 0.35 \end{pmatrix} \mathbf{y}_{t-1} + \begin{pmatrix} -0.40 & 0.25 \\ 0.10 & -0.25 \end{pmatrix} \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t, \tag{4.25}$$

with eigenvalues $\lambda' = (1.00 \; 0.79 \; -0.23 \; -0.41)$, which clearly show that the system has one unit root.

The linear restrictions needed to get the univariate restricted forecast, and the one-stage and two-stage multivariate restricted forecasts are, respectively

$$r = 0 \text{ with } \mathbf{C}_0 = \begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix},$$

$$r = 0 \text{ with } \mathbf{C}_1 = \begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\beta} \end{pmatrix},$$

and

$$\mathbf{R} = \widehat{\mathbf{Z}}_F \text{ with } \Sigma_{\mathbf{v}} = \text{MSE}\left(\widehat{\mathbf{Z}}_F\right) \text{ and } \mathbf{C}_2 = I \otimes \boldsymbol{\beta},$$

where $\mathbf{C}_0$, $\mathbf{C}_1$ and $\mathbf{C}_2$ are $1 \times H$, $1 \times 2H$ and $H \times 2H$ matrices respectively.

The Monte Carlo experiment was done with the following algorithm. Program routines were written in Matlab 6.5-Release 13 (MathWorks, Inc. Software).

Given T, H, p and the restriction $z_{T+H} = 0$:

1. *Generate a series* $\{\mathbf{y}_t\}_{t=-p+1}^{T+H}$ *with* (4.25). *Following the partition used by Lütkepohl (1991) we have a time series of length* p $\mathbf{y}_1, \ldots, \mathbf{y}_T$ *of length* T, *a presample* $\mathbf{y}_{-p+1}, \ldots, \mathbf{y}_0$ *and* H *future values of the time series,* $\mathbf{y}_{T+1}, \ldots, \mathbf{y}_{T+H}$.

2. *Estimate the* VEC *model given by (*4.24*) and compute the error correction term* $z_t = \widehat{\boldsymbol{\beta}}\mathbf{y}_t$.

3. *Compute the unrestricted forecast* $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})$ *and the trace of each MSE measure defined above.*

4. *Compute the one-stage restricted forecast* $\widehat{\mathbf{Y}}_{F,1}$ *and the trace of each MSE measure defined above.*

5. *For* z_t, *estimate the* ARMA(p,q) *model for* $p, q = 0, 1, 2$ *and* 3, *which minimizes the AIC criterion and compute the univariate restricted forecast as well as its MSE.*

6. *Compute the two-stage restricted forecast* $\widehat{\mathbf{Y}}_{F,2}$ *and the trace of each MSE measure defined above.*

Steps 1-6 were replicated $N = 1000$ times for the forecast horizons $H = 8, 16, 32$ and 64.

Tables 4.3 and 4.4 show the precision measures for the unrestricted forecasts, as well as for the one-stage and two-stage restricted forecasts. The precision measures for the model-based MSE and for the standard MSE confirm that expression (4.20) holds true. However, the precision measure results using the MSE of cointegrating combinations and the MSE of the triangular system in the short-run ($H = 8$) indicated no gain of $\widehat{\mathbf{Y}}_{F,2}$ over $\mathrm{E}(\mathbf{Y}_F|\mathbf{Y})$. However, for a forecast horizon large enough ($H \geq 16$) the results corroborated what inequality (4.20) says.

Table 4.3: Forecast precision measures: model-based and standard MSEs

H	$E(H)$* with model-based MSE			$E(H)$* with standard MSE		
	Unrestricted	1S restricted	2S restricted	Unrestricted	1S restricted	2S restricted
8	0.436	0.404	0.295	0.468	0.464	0.471
16	1.079	1.051	0.809	1.161	1.153	1.153
32	2.544	2.529	2.155	2.806	2.796	2.760
64	5.593	5.585	5.125	6.443	6.432	6.329

* See expression (4.21) for the definition of $E(H)$.

To gain some insight into this study, let us follow one realization of the algorithm. First, we generate a time series with equation (4.25) and estimate the

Table 4.4: Forecast precision measures that take contegration into account

H	$E(H)$* with MSE of cointegrating combinations			$E(H)$* with MSE of the triangular system		
	Unrestricted	1S restricted	2S restricted	Unrestricted	1S restricted	2S restricted
8	0.866	0.861	0.867	0.891	0.887	0.894
16	1.274	1.265	1.265	1.308	1.299	1.301
32	1.844	1.837	1.832	1.891	1.885	1.881
64	2.627	2.623	2.620	2.693	2.690	2.687

* See expression (4.21) for the definition of $E(H)$.

VEC of equation (4.24), so we have

$$\widehat{\Lambda} = \begin{pmatrix} \underset{(-0.13)}{-0.0005} \\ \\ \underset{(-1.95)}{-0.0147} \end{pmatrix}, \ \widehat{\boldsymbol{\alpha}} = \begin{pmatrix} \underset{(11.12)}{0.197} & \underset{(2.84)}{0.106} \end{pmatrix}', \ \widehat{\boldsymbol{\beta}} = \begin{pmatrix} 1 & -4.033 \end{pmatrix}$$

$$\widehat{\Pi}_1^* = \begin{pmatrix} \underset{(-3.39)}{-0.376} & \underset{(3.68)}{0.269} \\ \\ \underset{(0.64)}{0.149} & \underset{(-1.53)}{-0.236} \end{pmatrix}, \ \widehat{\sigma}_{\varepsilon_1} = 0.029 \text{ and } \widehat{\sigma}_{\varepsilon_2} = 0.062.$$

with $R^2 = 0.71$ and $R^2 = 0.34$ for y_1 and y_2 respectively. Figure 4.6 shows the observed and estimated series of this realization. The corresponding error correction term was fitted by an AR(2) model

$$(1 - \underset{(5.13)}{0.514}B - \underset{(2.55)}{0.257}B^2)z_t = \varepsilon_t; \qquad \widehat{\sigma}_\varepsilon = 0.250.$$

Figure 4.7 shows the estimated error correction term, its fitted values and the univariate restricted forecasts for $H = 16$. We chose this value of H by looking for the forecast horizon at which the error correction term becomes practically equal to zero, without forcing it to be zero by way of the restriction.

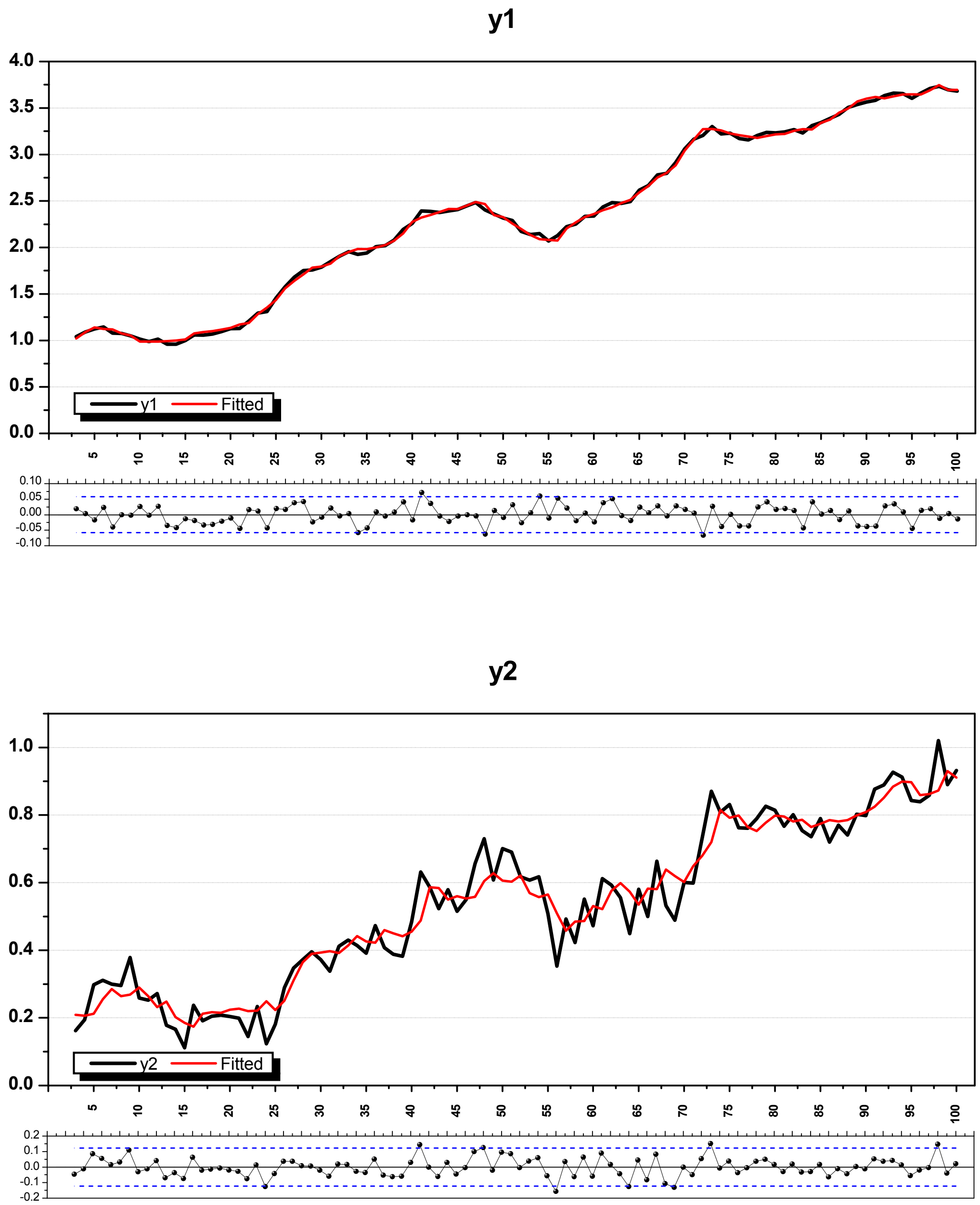


Figure 4.6: Cointegrated system with fitted time series, and residuals with horizontal lines correspondig to $\pm 2\widehat{\sigma}$

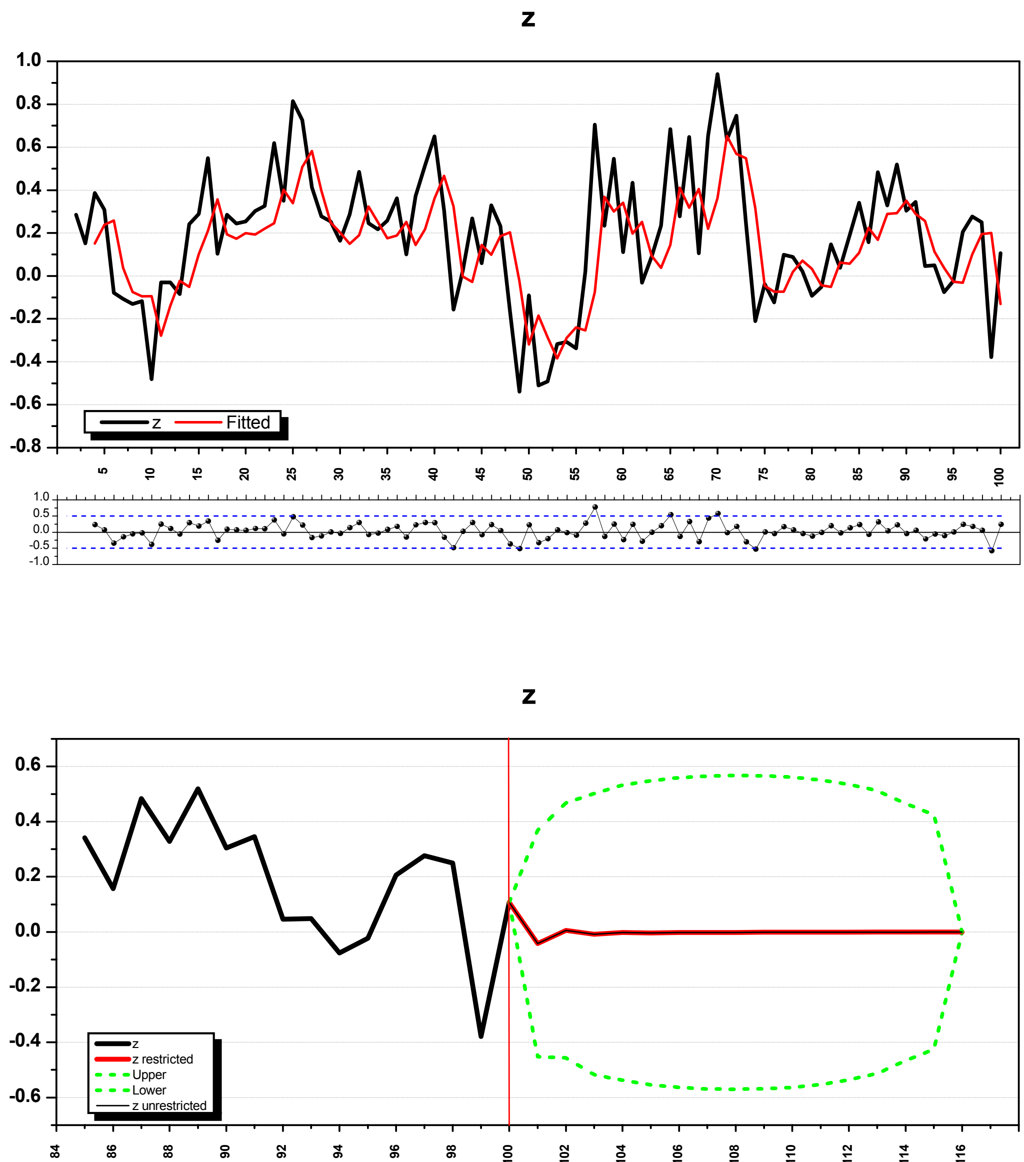


Figure 4.7: Estimated error correction term, fitted series and residuals with horizontal lines corresponding to $\pm 2\widehat{\sigma}$, at the top. Univariate restricted forecasts for $H = 16$, at the bottom

Figure 4.8 allows us to appreciate visually the fact that the precision of the two-stage forecast is higher than that of the one-stage forecast. Thus, if we do not incorporate the cointegrating relationship as a future contemporaneous constraint we may be wasting information.

4.6 Conclusions

We presented a forecasting methodology for multivariate time series that satisfy a contemporaneous binding constraint for which there exists a future target. Two ways of computing the forecasts are derived. The first one introduces the target as a linear restriction of the future values of the system, while the other introduces the target in the forecast of the aggregated variable (the contemporaneous binding constraint of the time series vector) which, in turn, is introduced as a restriction for the system forecasts. It was shown that the second way provides more precise forecasts than the first one, while the forecasts produced with the first way are more precise than the unrestricted forecasts. The methodology was illustrated with the income-expenditure system of the balance of payments account for the Mexican economy (with quarterly data). Here, the forecasts of the system are usually restricted to satisfy a government target for the deficit.

Since a cointegration relationship can be viewed as an unbinding contemporaneous constraint, the proposed methodology has immediate implications when forecasting cointegrated systems. Here we considered in detail the case of one unit

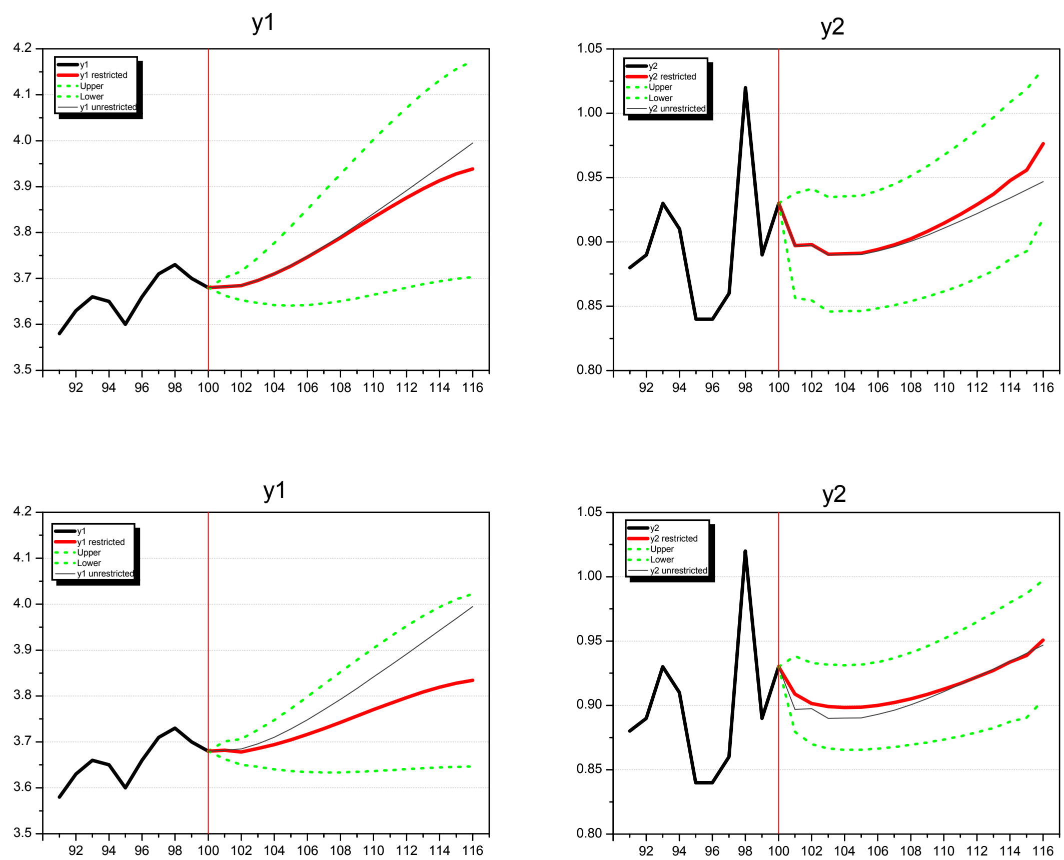


Figure 4.8: One-stage restricted forecasts with 50% probability intervals (with origin at 100) at the top. Two-stage restricted forecasts with 50% probability intervals, at the bottom

root. We started by requiring the equilibrium of the system to be reached at some finite point in the future, such a restriction was imposed by making zero the error correction term at that point. A Monte Carlo simulation of a Vector Error Correction (VEC) model with one unit root was done to compare the behavior of the unrestricted forecasts against those of the cointegrating restricted forecasts. Since the standard MSE has been severely criticized as a precision measure by several researchers, in our comparisons we used three additional precision measures.

BIBLIOGRAPHY

[1] Christoffersen PF, Diebold FX. 1998. Cointegration and long-horizon forecasting. *Journal of Business & Economic Statistics* **16**: 450-458.

[2] Clements MP, Hendry DF. 1993. On the limitations of comparing Mean Square Forecast Errors. *Journal of Forecasting* **12**: 617-637.

[3] Guerrero VM. 1989. Optimal conditional ARIMA forecast. *Journal of Forecasting* **8**: 215-229.

[4] Guerrero VM, Peña D. 2003. Combining multiple time series predictors: a useful inferential procedure. *Journal of Statistical Planning and Inference*. **116**: 249-276.

[5] Harvey A. 1993. *Time Series Models*, Second Edition. The MIT Press Cambridge, Massachusetts.

[6] Johansen S. 1988. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press.

[7] Lin JL, Tsay RS. 1996. Co-integration constraint and forecasting: an empirical examination. *Journal of Applied Econometrics* **11**: 519-538.

[8] Lütkepohl H. 1991. *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

[9] Pandher GS. 2002. Forecasting multivariate time series with linear restrictions using constrained structural state-space models. *Journal of Forecasting* **21**: 281-300.

[10] Pankratz A. 1989. Time series forecast and extra-model information. *Journal of Forecasting* **8**: 75-83.

[11] SHCP. 2004. *Criterios Generales de Política Económica 2005*. México: Secretaría de Hacienda y Crédito Público.

[12] Trabelsi A, Hillmer SC. 1989. A benchmarking approach to forecast combination. *Journal of Business Economic Statistics* **7**: 353 - 362.

[13] van der Knoop HS. 1987. Conditional forecasting with a multivariate time series model. *Economics Letters* **22**: 233 - 236.