MACROECONOMIC FUNDAMENTALS AND EMERGING MARKET ASSET PRICES

A Thesis

by

Osman Yilmaz

Submitted to the
The Graduate School Of Business
In Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

in the
Department of Business Administration

Özyeğin University
June 2016

To my little daughter who gave the word 'Life' a much richer meaning...

ABSTRACT

This thesis consists of three chapters which make empirical contributions to the field of emerging markets fixed income, real estate and financial markets. First chapter entitled *'Macroeconomics Fundamentals and Emerging Market Local Currency Debt'* focus on Emerging market (EM) local currency debt market which is largely absent from the academic literature, despite the increasingly important role of local currency debt for EM sovereign issuers and its increasing share in the portfolio of foreign investors. In this chapter, I investigate the effects of macroeconomic fundamentals on EM local currency bond markets using a dynamic factor approach based on a large panel of economic and financial time series. I find strong predictable variation in the EM local currency excess bond returns that is associated with macroeconomic activity. I provide evidence that the main predictor variables are the factors based on real economic activity that are highly correlated with measures of industrial and manufacturing production, but factors based on global financial factors also contain information about the future local currency bond returns. The predictive power of the extracted factors is not just statistically significant but also economically important. In the second chapter entitled *'Predictability of Emerging Market Real Estate Prices'* I approximate large information set of EM real estate market by large panel of economic and financial time series used in the first chapter. One of the main contributions of this chapter to the empirical literature is to document the mutuality of top three factors predicting the real house price fluctuations in a sample of leading emerging economies including Brazil, Mexico, South Africa, and Turkey. As two-thirds of the almost 50 systemic banking crises in recent decades were preceded by boom-bust patterns in house prices, I believe that my findings have important implications for policymakers and pension fund managers. Finally, third chapter entitled *'Forecasting Turkish Real GDP Using Targeted Predictors'* examines whether there is any merit of selecting a limited number of variables for superior forecasting performance. A number of recent studies in current literature discuss the usefulness of factor models in the context of GDP forecasting using large panels of macroeconomic variables. However, there is no consensus on how to identify informative variables from a large set of relevant indicators for the purpose of GDP prediction. Including too many variables in the analysis is likely to cause complications in extracting appropriate signal for the factor model framework. I empirically compare the forecasting performance of the dynamic factor model on various samples based on different selection criteria including my own. The forecasting exercise is performed for Turkish real GDP growth. My results show that the new sampling technique performs best as it attains first place in ranking for all backcast, nowcast and one-quarter ahead forecast periods.

ÖZETÇE

Bu tez gelişmekte olan ülkelerde (EM), sabit getirili menkul kıymetler, gayrimenkul ve finansal piyasalar alanlarında deneysel katkılar içeren üç bölümden oluşmaktadır. Makroekonomik Temeller ve Gelişmekte Olan ülkelerde Yerel Para Cinsi Tahviller başlıklı birinci bölümde, gelişmekte olan ülkelerdeki ihraççılar ve yabancı yatırımcılar açısından önemi artan, ancak akademik literatürde kendine yeteri kadar yer bulamayan yerel para cinsinden borçlanma piyasalarına odaklanılmaktadır. Bu bölümde, ekonomik ve finansal zaman serilerinden oluşan büyük bir panele dayalı dinamik faktör yaklaşımı kullanılarak, makroekonomik temel gerçekleşmelerin EM yerel para cinsinden tahvil piyasalarına etkisi araştırılmaktadır. Yine bu bölümde, EM yerel para cinsi tahvillerindeki aşırı getirilerdeki değişimin, makroekonomik gerçekleşmeler ile ilişkili ve tahmin edilebilir olduğu gösterilmektedir. Ayrıca, imalat ve sanayi üretimi gibi reel ekonomik aktiviteye dayalı faktörlerin ana belirleyici değişkenler olduğuna ve küresel finansal piyasalara ilişkin faktörlerin de tahvil getirileri hakkında tahmin edilebilir bilgi içerdiğine dair kanıtlar sunulmaktadır. Cıkarılan faktörlerinin aşırı getirileri tahmin etme gücü sadece istatistiksel olarak anlamlı değil, aynı zamanda ekonomik açıdan da önemlidir.

"Gelişmekte Olan Piyasalarda Gayrimenkul Fiyatlarının öngörülebilirliği" başlıklı ikinci bölümde ise ilk bölümde kullanılan ekonomik ve finansal zaman serileri kullanılarak EM gayrimenkul piyasaları öngörülmeye çalışılmaktadır. Bu bölümün deneysel literatüre temel katkılarından biri; Brezilya, Meksika, Güney Afrika ve Türkiye gibi önde gelen gelişmekte olan ülkelerin enflasyondan arındırılmış konut fiyatı dalgalanmalarını üç faktörle açıklayabilmesidir. Yakın geçmişte yaşanan yaklaşık 50 sistemik bankacılık krizinin üçte ikisinin konut fiyatlarındaki ani yükseliş ve sert düşüş

hareketleri sonucunda yaşandığı düşünüldüğünde, bu bölümdeki bulguların politika yapıcıları ve fon yöneticileri açısından önemli sonuçları olduğu ortaya çıkmaktadır.

Son olarak, Hedefli Parametreler Kullanarak Türkiye Reel GSYH Tahmini başlıklı üçüncü bölümde, üstün tahmin performansı elde etmek için kullanılan değişkenlerin sayısında sınırlamaya gitmenin önemi olup olmadığını incelenmektedir. Mevcut literatürde yer alan son çalışmalarda, bir dizi makroekonomik değişkenden oluşan büyük panelleri kullanarak GSYH tahmini yapan faktör modellerinin yararlılığı tartışılmaktadır. Ancak, GSYH tahmini amacıyla oluşturulan büyük gösterge setlerinin, bilgilendirici değişkenleri belirlemek için nasıl kullanılacağına ilişkin bir fikir birliği yoktur. Yapılan analizlerde çok sayıda değişken kullanmanın faktör modeli çerçevesinde uygun sinyalleri oluşturmada sorun çıkarması olasıdır.

Bu bölümde kendi modelim de dahil olmak üzere, farklı değişken seçimi kriterlerine göre oluşturulmuş dinamik faktör modellerinin tahmin performansını karşılaştırıyorum. Tahmin çalışmasını Türkiye reel GSYH üzerinde gerçekleştiriyorum. Sonuçlarıma göre, kullandığımız yeni örnekleme tekniği bir sonraki çeyrek GSYH değeri için tüm tahmin ve doğrulama analizlerinde ilk sırada yer almakta ve en iyi tahmin performansını göstermektedir.

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Professor Emrah Sener for his continuous support during the Ph.D. programme. He was always available to listen and to steer me in the right direction. In the course of time, our advisor-student relationship developed into a friendship which I deem quite precious. I also say thank you to Professor Vedat Akgiray , who as my co-advisor dedicated his precious time to guide my academic studies not also during Ph.D. programme but also during my master study at Bosphorus University. Besides my advisors, I am also grateful to the rest of my thesis committee.

A special thanks goes to, Sayad Baronyan, who has been a fellow companion during our Ph.D. journey as well as being a good friend and also a trustworthy colleague in the workplace.

Last, but not least, I thank my parents, for their altruism and for believing in me. My wife Can, for her unconditional support and for the weekends I had to spend at school instead of spending with her. And to everyone, who dedicated their lives to serving to prosper our beautiful country.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

MACROECONOMIC FUNDAMENTALS AND EMERGING MARKET LOCAL CURRENCY DEBT

1.1 Introduction

Most recent empirical analysis of emerging market bond market have been confined to frameworks in which researchers are implicitly assumed to exploit only a limited amount of information, despite the fact that EM central banks publish literally hundreds of economic and financial series. This chapter explores the feasibility of incorporating richer information sets into analysis of emerging market local currency debt market and investigate the extent to which bond returns are time-varying, how they are related to specific macroeconomic factors. I present a robust evidence that emerging market macroeconomic factors, especially the local real economic activity factor, has strong predictive power for excess bond returns even in the presence of financial predictors.

Emerging Market (EM) debt originated as a foreign currency (FC) denominated external debt market but over the last decade local debt (LC) has become a firmly established strategic asset class and it is now the LC issuance that dominates the sovereign debt marketplace.[1] LC debt offers higher returns than FC debt, or, equivalently, they attract lower prices because EM governments can default on their local currency debt, therefore their borrowing costs reflect both currency and credit risks. Evolution of local currency yield curves has been primarily from short-maturity issuance to longer-dated issuance mainly due to: (i) increase in creditworthiness due to

[1] The local currency government bonds are defined as bonds issued by the domestic government and denominated in local currency.

efforts by EM governments to cut their debt levels, stockpile foreign currency reserves and institute robust anti-inflationary measures and (ii) greater financial stability in EMs lead to development of most likely buyers of local currency debt such as local pension systems, insurance companies, and mutual funds and (iii) higher yields of LC debt relative to the traditional fixed income universe and their attractive risk and return attributes has made them a compelling component of global portfolios. Global investors seeking greater yield and diversification in their portfolios increasingly turn to emerging markets local debt. They invest more than 2.6 trillion dollar into the asset class and hold almost 32% of the total outstanding EM local currency debt as of December 2012. This represents a sizable pick-up from around 7% in mid-2005.

The nature of a LC bond premia risk is quite important for EM countries because it may affect both their ability to access international debt markets and the risk premium it must then pay to obtain capital. Furthermore, understanding the nature of LC risk profile is of first order importance for global fund managers. If LC debt is driven primarily by country specific factors, then standard portfolio diversification methods are available to manage risks embedded in LC debt. On the other hand, if LC debt risk is driven primarily by global risk factors, then there are major implications for the optimal allocation of investment capital across emerging market countries. Thus movements in LC debt markets have significant impact on local household welfare but also on the global financial stability. How do financial markets value local debts of emerging markets? Should global investor view LC debt market as a separate asset class ? How much of the yield stems from country's own fundamentals? Answers to these questions not only has a profound effect on the valuation of emerging market assets but also on important economic issues such as the cost of capital, international diversification benefits, and international risk sharing.

To examine these questions, I use two data sets that provide cross-sectional information on sovereign emerging economies. The first data set provides zero coupon yields of local currency bonds with maturities up to five years, with one year increments. The second data set provides information on economic and financial series at country level. My data set covers the time period between June 2006 and March 2014. The advantage of using this longer sample period for LC debt market is that it covers two major financial crises (US subprime crises and European sovereign debt crises) rather than just the relatively uneventful mid-decade period. I select the four major emerging sovereign bond markets (by notional amount outstanding) denominated in their local currency namely; Brazil, Mexico, South Africa and Turkey. These countries share three important features: (i) they belong to the J.P. Morgan EM-GBI index, an investable index for emerging market LC bonds, and (ii) they have large and liquid LC bond markets in which search and trading costs are low, and (iii) they offer long term LC bonds.

I proceed in the following steps. First, I apply dynamic factor analysis to summarize a large amount of macroeconomic and financial series in a relatively small number of estimated factors without having degrees of freedom problems. Essentially, this methodology enables me to condition the emerging market local currency bond price forecasts on a large information set involving more than 100 economic and financial variables. This is in sharp contrast to previous studies on EM bond market where only a few observed variables are included in the predictor set. This methodology has also advantage of accommodating data of different vintages and frequencies which is quite important because important economic and financial variables do in fact arrive at a variety of frequencies, including quarterly (e.g., GDP), monthly (e.g., capacity utilization production), weekly (e.g., unemployment), and continuously (e.g., financial asset prices). Second, I regress each estimated factor on each single variable included in the data set to assign economic interpretation to factors. This allow us to see whether our estimated factors represent the common business cycle of the original EM economic and financial series. Finally, I run a forecasting experiment on LC bond returns by conditioning on a rich information set instead of only a few variables. Although the estimates of the predictable dynamics in excess local currency bond returns depend on the estimated factors, I use a statistical criterion for choosing parsimonious models of relevant factors which makes our forecasting experience less dependent on a handful of predetermined variables. I choose among a range of possible linear and nonlinear specifications for the forecasting regressions for two, three, four, and five year LC excess bond returns for each country and examine the forecasting performance for the short and long horizons.

I report a number of novel empirical results. First, I show that eight to fifteen common factors account for about 40% to 70% of the variation in the hundreds of economic series. My statistical criterion point us to five factors for Brazil and Mexico, six factors for Turkey and eight factors for South Africa that have important additional forecasting power for LC excess bond returns. I show that strong predictable variation in LC excess bond returns that is associated with EM macroeconomic activity and provide the evidence that two key factors: *local real economic activity factor* and *global financial* factor have the predictive power that is not just statistically significant but also economically important. I also find that factors associated with local real economic activity have the most significant predictive power for excess bond returns. This point us to the fact that on average, changes in local currency bond risk premia appear to be more closely related to sovereign own local economy than changes in the global factors. Second, I find that the adjusted R^2s for the forecasting regressions are intriguing. In general, these R^2s are fairly high for all countries, indicating that the estimated factors capture much of the variation in LC bond return. These factors are statistically significant predictors of bond returns and explain 24%, 31%, 44%, and 49% of the variation one year ahead in the two-year return for Mexico, Brazil, Turkey and South Africa respectively. Interestingly, I observe that forecasting power increases as we move along the LC yield curve and these factors explain up to 50%, 51%, and 64% of next year's excess return on the three, four, and five-year bonds respectively. Finally, to shed light on the underlying nature of these key identified factors, I also examine the cross-country correlation structure of the estimated factors. My findings point to the fact that the existence of a unique risk premium (via local real economic activity) in local currency debt returns validates the view that LC debt markets as a separate asset class.

I believe that the findings in this chapter have very important implications for policymakers and pension fund managers. I identify the key factors explaining variation in local risk premia across major emerging countries and over time. Policymakers may find it beneficial to address such risk factors in order to reduce their local and foreign borrowing costs as well as the probability of a crisis. In particular, EM governments can try to explicitly hedge exposures to macroeconomic risks more effectively by using the innovative financial contracts in international financial markets ([1] and [2]). At present most of the developed market pension schemes allowed to have some exposure to emerging markets as long as they are rated investment grade and above by credit rating agencies. In fact, investment-grade EM countries represented in the JP Morgan Emerging Markets Bond Index (EMBI) Global Diversified Index jumped from less than 2% in 1994 to 63% in 2013. Thus developed markets pension fund managers may find my findings to add value in their asset allocation efforts which should further enhance diversification of pension assets and potentially reduce total funding level risk.

1.2 Related Literature

This chapter is related to research in macro-finance and empirical asset pricing that investigate the possible empirical linkages between macroeconomic variables and EM asset returns. A number of macroeconomic variables has been proposed in the literature to provide a direct line linking asset return predictability to economic fundamentals such as investment/capital ratio ([3]), growth of non-farm payroll employment ([4]), ratio of labor income to total income ([5]), consumption to wealth ratio ([6]) and output gap ([7]). Most of the EM debt market research has been confined to FC denominated bonds and developed in two directions. The first deals with empirical determinants of FC sovereign spreads and establish several explanatory variables, both global and country specific. Examples of this line of research includes papers that examine the FC denominated bonds empirically as a function of a bulk of country specific solvency variables such as reserves/GDP ([8]), export growth ([9]), terms of trade ([10]), and also include papers that concentrate on global factors such as world interest rate ([11]), the U.S. high yield spread ([12]), global liquidity ([13]). The second deals with theoretical pricing models that relate FC spread to VIX, political factors and oil prices, but macroeconomic fundamentals do not directly enter the estimation of the pricing models ([14] and [15]). Despite the growing body of theoretical and empirical works rationalizing economic variables and bond risk premia in EM foreign currency debt markets, there is little direct research on local currency debt markets. Indeed, [16] summarize the state of asset valuation in emerging markets and emphasize that EMs provide a challenge to existing models and beg the creation of new models. Table 1 is a summary of key literature on asset pricing of both developed and emerging markets assets. I contribute to the literature on emerging markets bond return forecastability by showing that emerging markets macroeconomic fundamentals have important predictive power for LC government bond markets. This chapter also adds to the literature to enhance our understanding of the economics of time-varying risk premiums in emerging markets.

The remainder of the chapter is organized as follows. Section 2 describes the data and the yield curve specification used to estimate the zero curve. Section 3 provides a general discussion of the dynamic factor model and its role in the forecasting framework. Section 4 discusses the main properties of the extracted factors and focus on their forecasting performance on local currency bond return. Some concluding remarks are summarized in the conclusion.

Table 1: Related Literature on Asset Pricing Theory

		Asset Prices and Macro Factors: Developed Markets			
Authors	**Region**	**Asset Pricing Model**	**Period**	**Methodology**	**Relationship**
Cochrane (1991)	United States	Investment/Stock Returns	1947-1987	Panel Regression	Positive
Piazessi-Swanson (2004)	United States	Employment/Fed Funds Rate	1988-2004	Regression	Negative
Menzly et al. (2004)	United States	Stock Prices/Dividends	1947-2001	General Equilibrium Model	Negative
Lettau-Ludvigson (2005)	19 OECD Countries	Consumption/Stock Returns	1948-1999	General Equilibrium Model	Positive
Copper-Priestley (2009)	G7 Countries	Output Gap/Stock Returns	1948-2005	Regression	Negative
		Asset Prices and Macro Factors: Emerging Markets			
Authors	**Region**	**Asset Pricing Model**	**Period**	**Methodology**	**Relationship**
Edwards (1986)	Brazil, Mexico	International Reserves/EM Sovereing Spreads	1980-1985	Regression	Negative
Cline et al. (1997)	12 Emerging Countries	Export Growth/EM Sovereign Spreads	1986-1997	Panel Regression	Negative
Hilscher-Nosbusch (2010)	31 Emerging Countries	Terms of Trade / EM Sovereign Spreads	1994-2007	Regression	Negative
Uribe-Yue (2003)	7 Emerging Countries	US Bond Yields/EM Sovereign Yields	1994-2001	VAR	Positive
Gonzalez Rozada-Yeyati (2008)	Emerging Countries	US High Yield Spread/EM Sovereign Spreads	1994-2005	Panel Error Correction Model	Positive
Peiris (2010)	10 Emerging Countries	Global Liquidity/EM Sovereign Spreads	2000-2009	Panel Regression	Negative

1.3 Data Description

The EM debt market is divided into two main categories: (1) foreign currency (external) debt, which is issued in a hard currency[2] other than that of the issuers domestic currency and (2) local currency (domestic) debt, which is denominated in the domestic currency of the issuer. My empirical investigation concentrates on sovereign local currency debt markets for four major emerging market governments namely; Brazil, Mexico, South Africa and Turkey. My choice of emerging market countries is mainly constrained by the lack of sufficient numbers of LC liquid and transparent bonds outstanding and available economic series. Furthermore, all four sample countries belong to the J.P. Morgan EM-GBI index, an investable index for emerging market LC bonds. I gather cross-country data on macroeconomic fundamentals and zero-coupon yield curve data from a variety of sources over the period from June 2006 to March 2014. The length of the sample period is constrained by the availability of long term local currency bond data.

Yield curve data: I collect daily zero-coupon yield data for our sample countries from two sources. First, I use zero-coupon LC curves constructed by the central bank of government agencies when they are available. Thus zero coupon yield curve data for Brazil and Turkey are collected from Brazilian Financial and Capital Market Associations (ANBIMA) and Central Bank of Turkey respectively. All bonds have fixed coupon rates and are not callable, puttable or convertible. I also exclude bonds governed by foreign jurisdictions. Second, when national data are unavailable, I utilize the Bloomberg Fair Value (BFV) curve. As discussed at [17] , BFV curves are par yield curves estimated by Bloomberg on actively traded bonds using piecewise linear zero-coupon curves and often serve as the benchmark reference rate in respective currencies. Thus zero coupon yield curves for Mexico and South Africa are collected

[2]Hard currencies refer to globally tradable, reliable, and stable currencies such as the US dollar, yen or euro.

from Bloomberg Fair Value par to zero. For curve estimation, standard [18] methodology is used to convert the par yield curves into zero curves. Constructed yields on zero-coupon bonds with maturities of one, two, three, four, and five years are used for my analysis.

Economic series: Constructing a rich database is a crucial step to extract the information in my analysis. Thus I estimate factors from a balanced panel of 124, 111, 108 and 118 economic series for Brazil, Mexico, South Africa and Turkey respectively. These series are selected to represent broad categories of economic series and include a large number of variables which allow us to kill the idiosyncratic variance over the cross-section. A two-step process is applied in order to adjust the data before the analysis. First of all, these series are transformed to stationary series because the estimation of my factor framework requires stationary time series. Secondly, these series are normalized by subtracting their mean and dividing by their standard deviation in order to have a zero sample mean and unit variance. This standardization is necessary to avoid over weighting of the series with large variance. Data were further regrouped into: *Real Economics Activity Variables*: Industrial production, retail sales, international trade and car sales. *Housing Variables*: House price index and real estate unit sold. *Labor Market Variables*: Employment and unemployment. *Prices*: Consumer prices, producer prices and commodity prices. *Money Credit Quantity Aggregates:* Monetary base, money supply (M1-M4) and deposits (time, demand and fx). *Financial Variables*: Exchange rates, interest rates and stock prices. These distinctions and the variety of economic and financial data series allow us to get a deeper insight into the dynamics of the LC debt markets. Notice also that my economic series include both pure macroeconomic and financial variables. This is important because fluctuations in the aggregate emerging market economy consist of substantial co-movement in financial and real economic variables. As I argue in section IV, one of my most important finding is that the estimated factors are highly correlated with both real economic activity and financial indicators. I include more detailed descriptions of the data sets and my data sources in the Data Appendix.

1.4 Theoretical Framework

In this section, I first present general overview of econometric and factor models. Then I explore the theoretical background of dynamic factor models and describe how to estimate the factors. Finally I discuss the various economic applications of factor models and describe how to run predictive regressions for LC debt market in which the predictor set includes time series of common factors from the factor analysis.

1.4.1 Econometric versus factor models

Researchers, policymakers and market practitioners nowadays have more economic and financial data series at a more disaggregated level at their disposal than ever before. To benefit from this large amount of information, one needs an appropriate model to create accurate forecasts, as well as to test economic theories. Incorporating more than a few variables under usual time series models such as autoregression and vector autoregression is not possible. Because a typical regression based model would run into a scarce degrees of freedom problem if the number of parameters to estimate is large with respect to the number of observations. Over last two decades, literature splits into two camps to solve the degrees of freedom problem. The first stream proposes to use variable selection procedures such as general-to-specific algorithm ([19]), regression shrinkage ([20]) and the simulated annealing (Kapetanios, (2007)). The problem with these proposed procedures is that the econometric models are still based only on the few chosen variables, and much of the information carried by the large data set would be lost. Second stream of the literature investigates to create an index model by using all the information available in the data set. This stream is developed in two directions. The first deals with principal component analysis

(PCA) and the second is the dynamic factor analysis (DFA)[3]. Factor models became more popular than many other econometric models for three fundamental reasons: (i) they can cope with many variables without running into scarce degrees of freedom problems (ii) they can remain agnostic about the structure of the economy and do not need to rely on overly tight assumptions, and (iii) they can lead to more precise forecasts which prevents policy makers from reacting to idiosyncratic movements.

1.4.2 Dynamic factor models and the estimation of factors

Since the first generation of factor models were introduced by [21] and [22] , dynamic factor models have been applied to many fields of macroeconomics and finance such as arbitrage pricing theory ([23]) and macroeconomics (Bernarke and Boivin (2003)). Similar to bond market applications of [24] , in my forecasting experiment, I implement the factor method of [25].[4] I consider a panel of observable economic variables $X_{i,t}$; where i denotes the cross-section unit $i = 1; ...; N$, and t refers to the time index $t = 1; ...; T$. I transform these series into stationary variables with zero mean and unit variance and label this data set as $x_{i,t}$. I specify the factor structure as,

$$x_{i,t} = \lambda_i f_t + \xi_{i,t} \tag{1}$$

where f_t is an $k \times 1$ vector of latent common factors, λ_i is a corresponding $k \times 1$ vector of latent factor loadings, and ξ is the idiosyncratic term.[5] The key implication of my factor framework is that the variation of each of the N economic variables

[3]While for classical principal component analysis, the number of economic time series (N) is relatively small but time periods (T) are large, for dynamic factor analysis both economic time series and time periods can be large and converge to infinity. By large, I mean large in the cross-section, for instance N = 100+ and T = 100+ depending on the frequency of the data.

[4]Given the sample size we use, this method also performs similarly well compared to more complex dynamic factor models of [26]

[5]It is even possible to introduce cross-sectional correlation among the idiosyncratic terms. This is ensured by imposing the condition that the contribution of the covariance of the idiosyncratic terms to the total covariance of x as N gets large is bounded (by a constant M) :

can be decomposed into a common component, f_k, that captures the cross-sectional co-movement and an idiosyncratic component, ξ and the key advantage of this framework is that it allows the data the maximal freedom to speak for themselves without any fundamental or economic structure imposed. I use non-parametric estimation approach based on principal components to estimate the factors.[6]. While [28] uses the time domain method, [29] suggests to use frequency domain non-parametric estimation technique. I prefer time domain method due to its simplicity and speed. Thus f is a $T \times r$ matrix of estimated values given by $\sqrt{T}$ multiplied by the r largest eigenvalues of the $T \times T$ matrix xx'/NT. Letting $\Lambda = (\lambda_1, \lambda_2, ..., \lambda_r)'$ denote the $N \times r$ matrix of factor loadings, the estimates of the factor loadings are then calculated by $\Lambda = x'f^T$.

Since my purpose is to forecast the LC bond return, I now move to define excess bond return calculation. I use the following notation for log bond prices $p_t^{(n)} =$ log price of n-year discount bond at time t. Then the log yield is $y_t^{(n)} = -\frac{1}{n}p_t^{(n)}$. I write the log holding period return from buying an n-year bond at time t and selling an n-1 year bond at time t+1 as $r_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)}$. I denote excess log returns by:

$$rx_{t+1}^{(n)} = r_{t+1}^{(n)} - y_t^{(1)} \tag{2}$$

To integrate the lag of the estimated factors and additional predictor variables (Z_t) at time t into our framework, dynamic factor model is constructed as,

---

$$N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}|E[\xi_{i,t}\xi_{j,t}]| \leq M$$

[6] A number of other estimation techniques has been proposed in a related literature such as Bayesian estimation technique ([27] and EM algorithm ([] Junbacker & Koopman (2008)).

$$rx_{t+1}^{(n)} = \beta(L)f_t + \gamma(L)Z_t + \xi_{i,t} \quad (3)$$

$$x_{i,t} = \lambda_i(L)f_t + \xi_{i,t} \quad (4)$$

for $i = 1, ..., N$, where $\xi_t = [\xi_{1,t}, ...\xi_{N,t}]$ is a N x 1 idiosyncratic term and $\lambda_i(L)$, $\beta(L)$ and $\gamma(L)$ are lag polynomials in non-negative powers of L. It is assumed that $E(\xi_{t+1}|f_t, Z_t, X_t, f_{t-1}, Z_{t-1}, X_{t-1}, ...) = 0$. [24] use [30] [30] forward rate factor (Zt) as a forecasting benchmark.

Determination of the number of factors representing the relevant information in the data set is a delicate issue. Various techniques have been proposed recently to determine the optimal number of factors in large panels. For instance, [31] and [25] suggest consistent selection procedures based on principal components, [32] propose an informal criterion based on the portion of explained variances and a test procedure based on the canonical correlation respectively.[7] I follow [34] to decide optimal number of factors.

1.4.3 Empirical applications of dynamics factor models

I identify three potential economic applications of dynamic factor models for EM economies; forecasting, monetary policy analysis and construction of economic activity indicators. I briefly explain each of them and focus on the first one in my empirical analysis.

(i) Forecasting: Factor models are widely used in central banks and research institutions to predict economic variables and analyze monetary policies in developed markets (see ([28], [35] for US, [36] and [37], [38] for the Euro zone and [39] for the

[7] Proposed methodologies result in substantial variations in the number of optimal factors. The results also show that the inclusion of a few additional variables may have a substantial effect on the number of factors. For example, for large US data set, while [25] find seven dynamic factor, [33] use a different methodology and identify four factors.

UK). Despite its popularity in developed markets, factor based analysis with EM economic and financial data sets is very limited.[28] and [40] show that forecasting performance of factor models are more successful compared to well known benchmark models such as autoregression and vector autoregression due to their effectiveness in reducing the dimension of the predictor set in which different sources of information shape each factor in the predictor set and their robustness to structural instability, which often plagues predictive regressions.

(ii) Analysis of monetary policy: Monetary policy-makers monitor literally thousands of data series from disparate sources and exploit only a limited amount of information to analyze the effect of their monetary policies. Literature explores the possibility of following policy rules to follow few critical variables. For example [41] recommended a policy rule in which movements in the interest rate can be traced to movements in macro variables and uses two macro variables; annual inflation rate and output gap. Recently literature also investigates the possibility of developing an expert system by utilizing dynamic factor models (see [42] and [43]) that could aggregate diverse information and provide benchmark policy settings.

(iii) Construction of economic activity indicators: Aggregate business conditions are of central importance for policy makers and market practitioners. The two most prominent examples that use dynamic factor model to create an index of economic activity, are the Chicago Fed National Activity Index, for US, and the EuroCOIN (see [44]), for the Euro Area.[8] [45] also show that it is possible to measure macroeconomic activity in real time with that dynamic factor model.

[8] Chicago Fed National Activity Index is simply the first static principal component of a large macro data set (see http://www.chicagofed.org/economic research and data/cfnai.cfm.). EuroCOIN is estimated as the common component of euro-area GDP based on dynamic principal component analysis (See http://www.cepr.org/data/eurocoin/).

1.5 *Empirical Analysis*

I divide the empirical analysis into three parts. In the first part, I extract a number of factors and provide an economic interpretation of these factors. In the second part, I identify possible linear and nonlinear specifications for the forecasting regressions and conduct a forecasting experiment on LC bond returns. Finally, following [30], I form a single linear combination of these estimated factors and assess its forecasting performance.

1.5.1 Extraction of the factors:

The number of factors f_t which are mutually orthogonal by construction is determined by the information criteria (IC). Figure 1 shows that, in each of these four major emerging markets, over the entire span of time we consider, only a handful of factors are needed to explain more 50% of the macroeconomics series co-movement. The IC indicate that the factor structure is well described by eight common factors for Brazil, South Africa and Turkey and fifteen common factors for Mexico. Moreover, the incremental power of each additional factor declines quite sharply for each country. For example first factor for Brazil explains 14% of the covariance, with the second factor explaining an incremental 13% or so, cumulating up to about 27%, and so on. By the time a seventh component is added, the amount of variance left to explain is less than 50%.

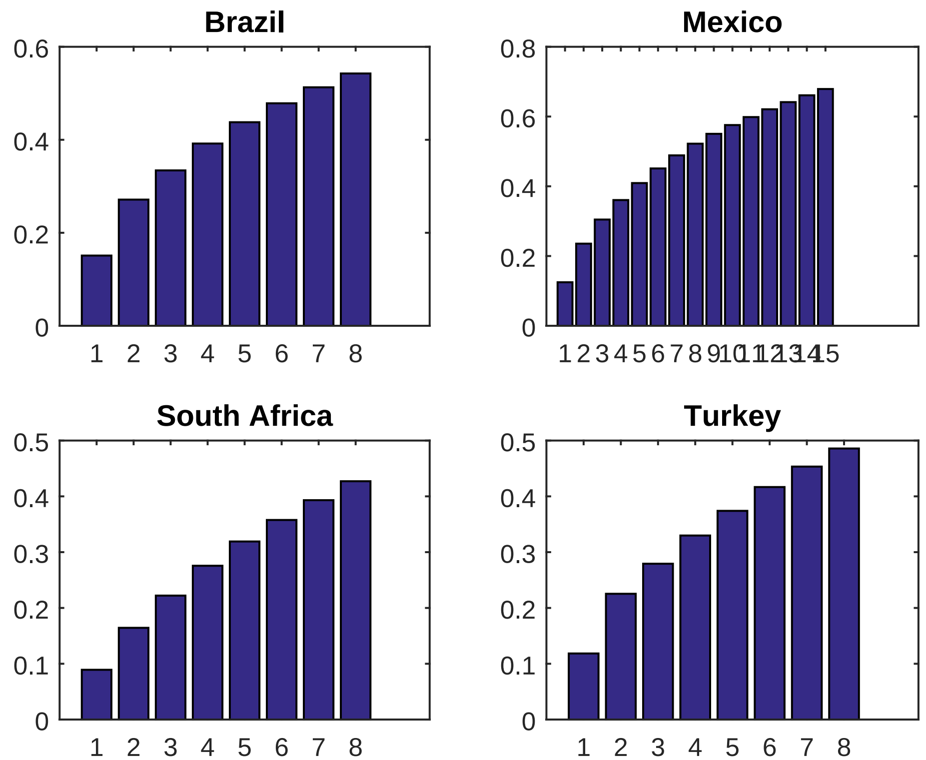


Figure 1: This figure shows overall variance contributions of factors selected by the IC criterion for the indicated countries.

I also investigate the specifications with lagged values of the estimated factors $\beta(L)f_t$ as shown at Eq(3) and find that additional lags contain little information for future local currency returns. From these estimated factors f_t, we form range of possible linear and nonlinear specifications for the forecasting regressions. I use the following form predictive regression in my analysis:

$$rx_{t+1}^{(n)} = \beta F_t + \epsilon_t \tag{5}$$

where $F_t \subset f_t$. Eq (7) is nested within the factor-augmented regression, generating a convenient framework to assess the importance of $x_{i,t}$ via F_t. The distinction between F_t and f_t is important, because factors that are pervasive for the panel of data $x_{i,t}$ do not need to be important for predicting $rx_{t+1}^{(n)}$. To determine the composition F_t, I form different subsets of f_t. For each candidate set of factors, F_t, I regress rx_{t+1} on F_t. I choose the the preferred set of factors F_t minimizing the Bayesian information criterion (BIC). The Table 2 shows factor specifications for each country.

Table 2: Factor specifications for each country

Country	**Selected Factors (BIC Criterion)**
Brazil	$F_{2,BR}$, $F_{3,BR}$, $F_{4,BR}$, $F_{5,BR}$, $F^2_{3,BR}$
Mexico	$F_{4,MX}$,$F_{5,MX}$,$F_{8,MX}$,$F_{11,MX}$,$F_{15,MX}$
South Africa	$F_{1,SA}$,$F_{2,SA}$,$F_{4,SA}$,$F_{5,SA}$,$F_{6,SA}$,$F_{8,SA}$,$F^2_{1,SA}$,$F^2_{5,SA}$
Turkey	$F_{1,TR}$,$F_{4,TR}$,$F_{6,TR}$,$F^2_{1,TR}$,$F^2_{2,TR}$,$F^2_{8,TR}$

While the use of dynamic factor analysis with IC criterion allows us to have a much larger set of predictor factors, the BIC criterion provides an efficient way of choosing among summary factors by indicating whether these variables have important additional forecasting power for excess bond returns. Figure 2 shows that BIC criterion selects five factors for Brazil including one nonlinear factor $F^2_{3,BR}$, five factors for Mexico, eight factors for South Africa including two nonlinear factors $F^2_{1,SA}$,$F^2_{5,SA}$ and six factors for Turkey including three nonlinear factors $F^2_{1,TR}$, $F^2_{2,TR}$,$F^2_{8,TR}$. These factor representations account for about 35%, 40%, 70%, and %45 the variation in the local currency excess return series of Brazil, Mexico, South Africa and Turkey respectively. These common statistical factors often have macroeconomic content primarily first and second factors can be related back to more structural and fundamental macroeconomic variables. Thus in my empirical analysis, I primarily focus on the first factor ($F_{\mathbf{4,BR}}$, $F_{\mathbf{5,MX}}$, $F_{\mathbf{1,SA}}$ and $F_{\mathbf{6,TR}}$) that explains the largest fraction of the total variation in excess bond returns rx_t for each country, where total variation is measured as the sum of the variances of the individual rx_t and the second factor ($F_{\mathbf{2,BR}}$, $F_{\mathbf{4,MX}}$, $F_{\mathbf{6,SA}}$ and $F_{\mathbf{4,TR}}$) that explains the second largest fraction of the total variation in excess bond returns for each country.

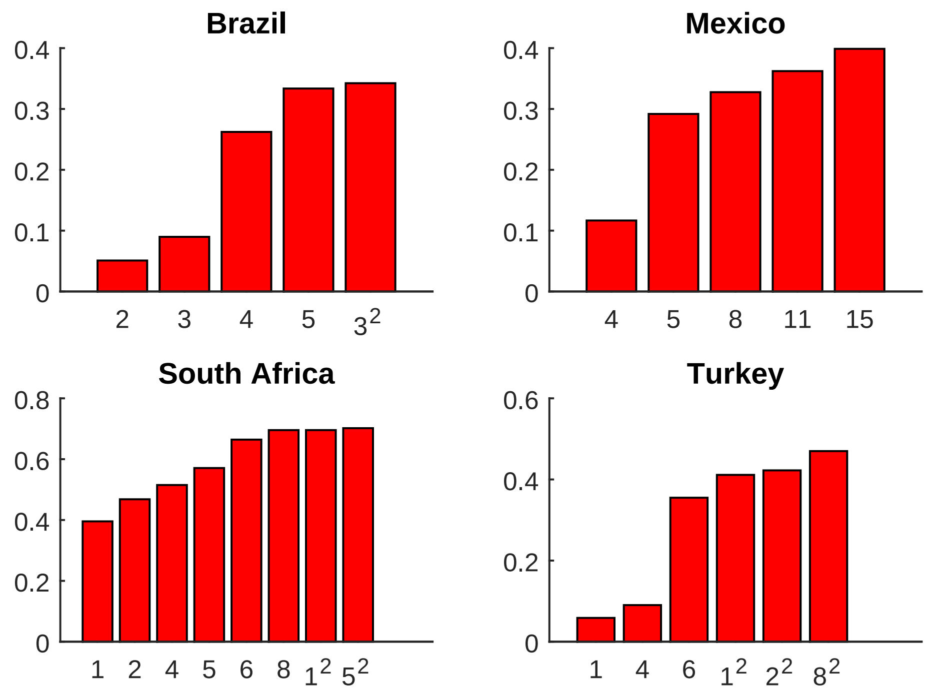


Figure 2: This figure shows the R^2 contributions of linear and non-linear factors selected by the BIC criterion for the indicated countries.

1.5.2 Economic interpretation of the estimated factors:

Having established that (i) only a small number of factors are needed to explain local currency bond returns and that (ii) many of these factors relate to common macroeconomic variables, now the daunting task ahead is to give these unnamed things some macroeconomic content. It is a critical issue to obtain a meaningful identification of the estimated factors. Because I try not only to find few common factors that explain a large economic series, containing numerous variables but also I am interested to see whether the estimated factors represent the common business cycle of the original variables.[9] To assign an interpretation to my estimated factors, I check the R^2 of the regression of each factor on each single variable included in the data set. Figures 3 to 10 show the R^2 statistic as bar charts from regressions of each of the individual series in my data set onto each estimated factor, one at a time for each country. The individual series that make up the data set are grouped by a broad category and labeled using the numbered ordering given in the Data Appendix. Figure 3 to 6 show that $F_{\mathbf{4},\mathbf{BR}}$, $F_{\mathbf{5},\mathbf{MX}}$, $F_{\mathbf{1},\mathbf{SA}}$ and $F_{\mathbf{6},\mathbf{TR}}$ that explains the largest fraction of the total variation in excess bond returns, load heavily on measures of real economic activity (e.g., industrial production and capacity utilization) for all countries but also on measures of housing variables for Brazil and South Africa and employment variables for Mexico and Turkey. This factor explains up to 62%, 40%, 67% and 48% of total variation for the real economic activity variables for Brazil, Mexico, South Africa and Turkey respectively. Figure 3 to 6 displays little correlation with prices and financial variables. However, Figure 7 to 10 show that $F_{\mathbf{2},\mathbf{BR}}$, $F_{\mathbf{4},\mathbf{MX}}$, $F_{\mathbf{6},\mathbf{SA}}$ and $F_{\mathbf{4},\mathbf{TR}}$ factors load heavily on measures of the aggregate financial variables (e.g., exchange rate and bond yield) and displays very little correlation with economic activity and employment variables. This factor explains up to 46%, 32%, 65% and 66% of the

[9] This is not a necessary step for my forecasting analysis. For example, [28] use dynamic factor analysis only to create a number of indexes to improve their macroeconomic forecasts.

total variation for the financial variables for Brazil, Mexico, South Africa and Turkey respectively. My findings point to two key factors namely real and financial variables that load heavily on macroeconomic variables might have substantial predictive power for local currency bond returns.

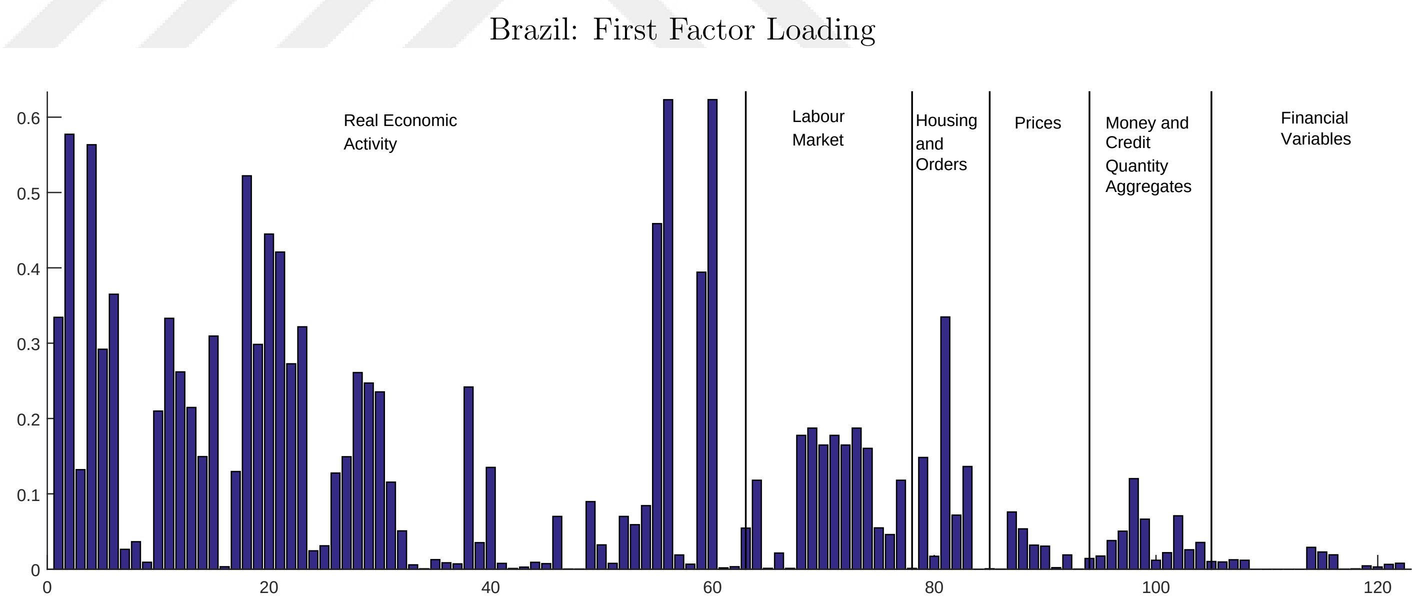


Figure 3: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{4,BR}}$ for Brazil. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

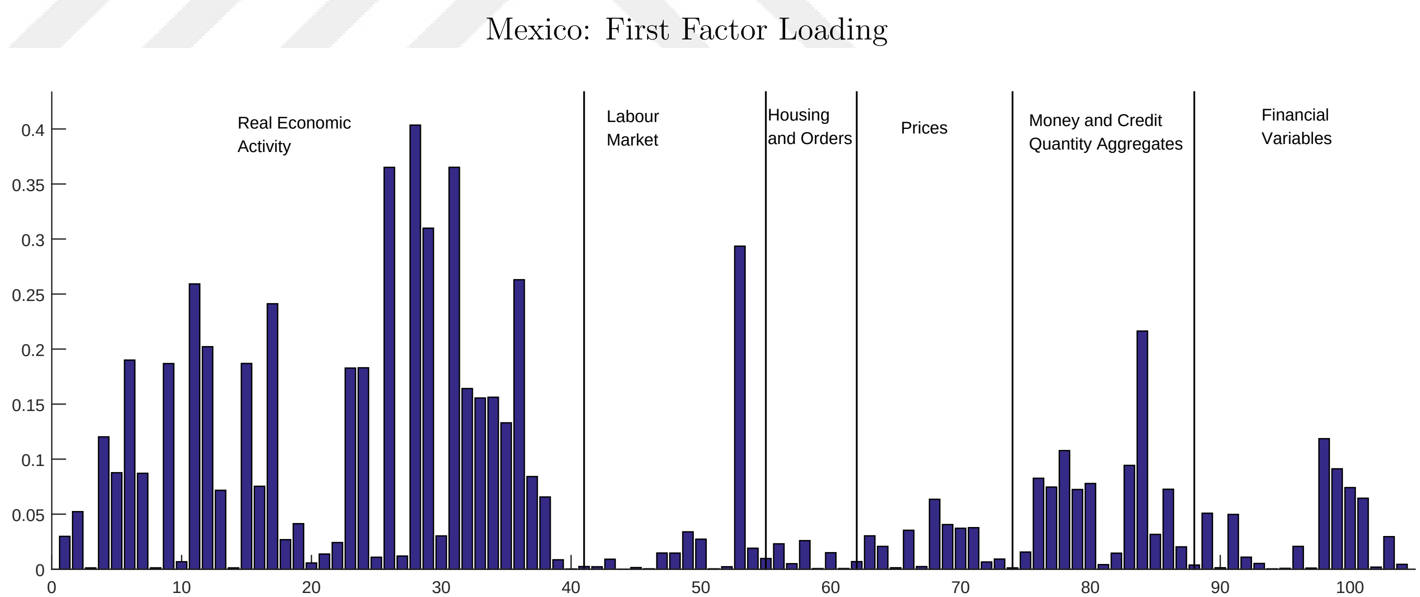


Figure 4: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto $F_{\mathbf{5,MX}}$ for Mexico. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

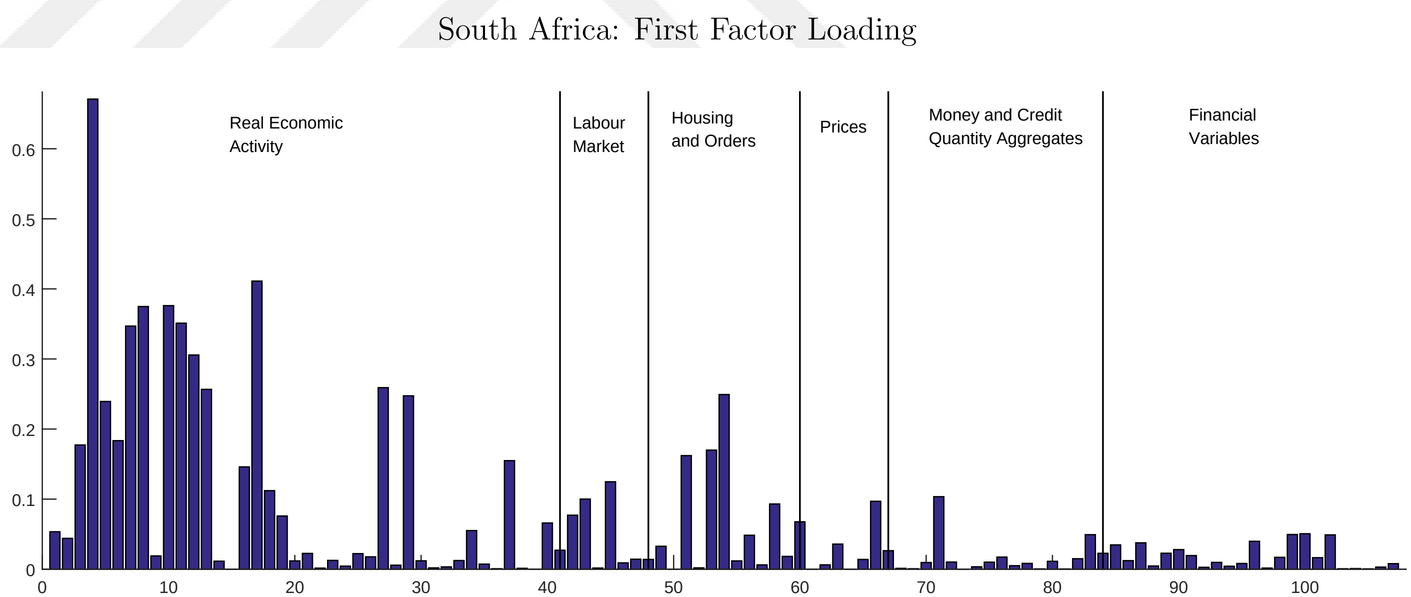


Figure 5: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{1},\mathbf{SA}}$ for South Africa.The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines for.

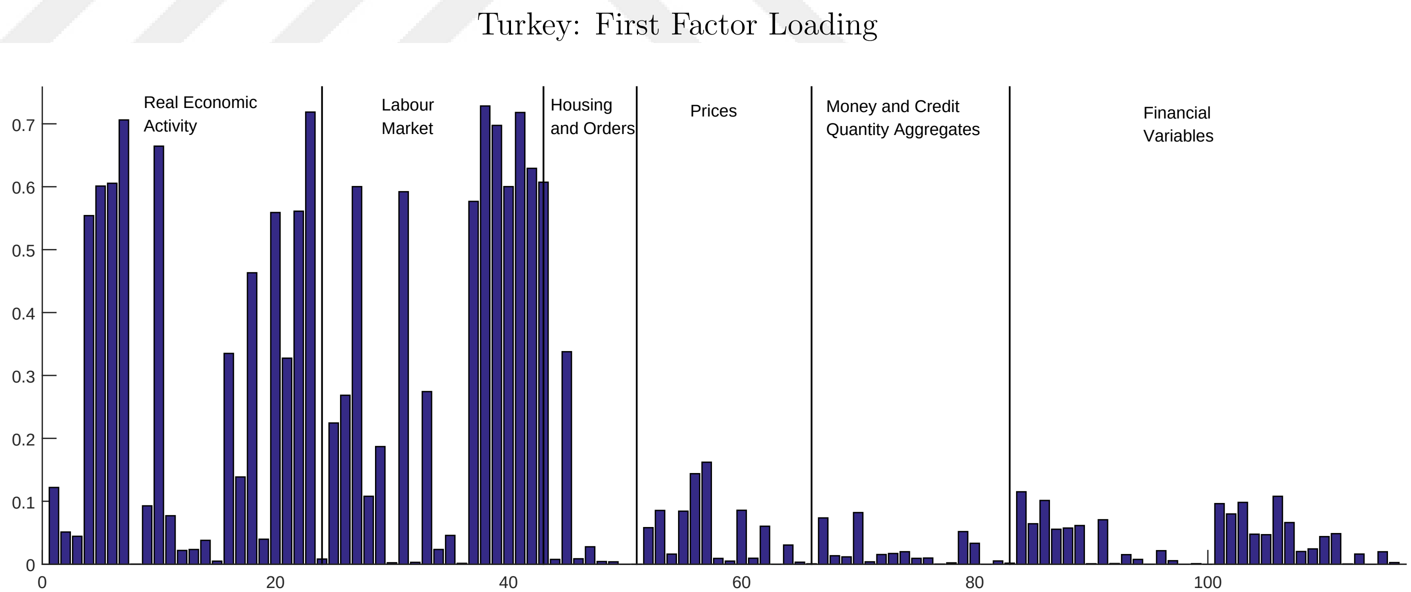


Figure 6: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto the first factor loading, $F_{\mathbf{6,TR}}$ for Turkey. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines for each country.

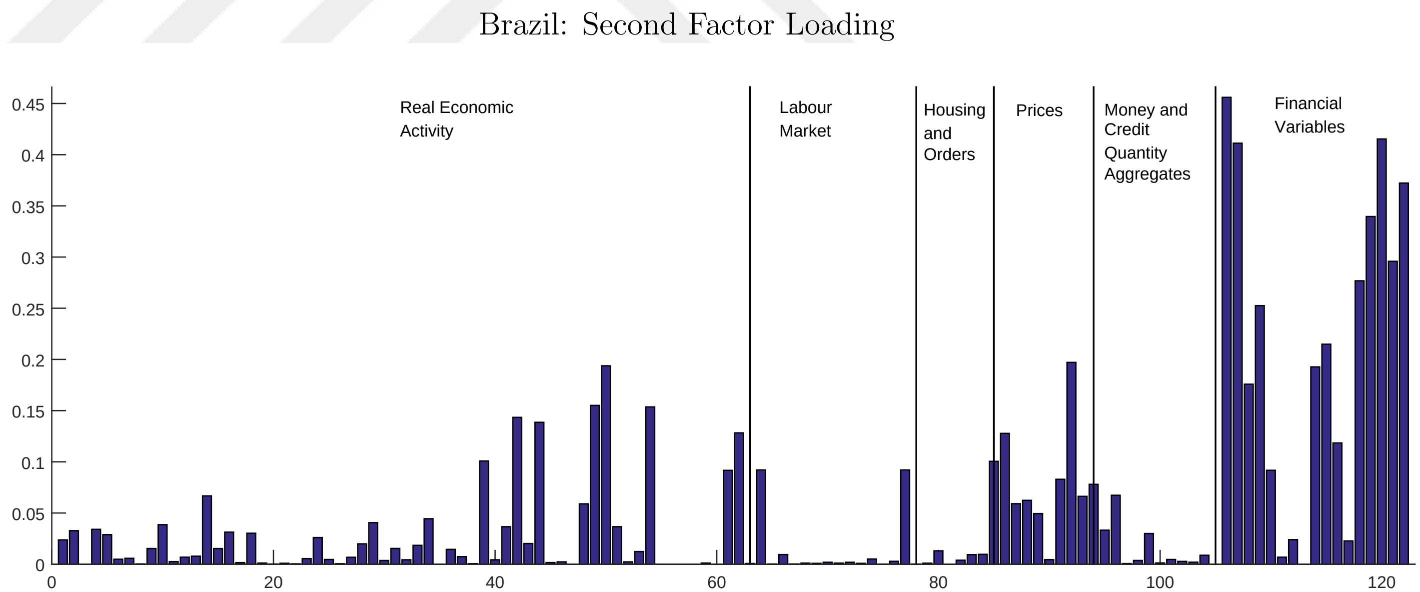


Figure 7: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{2,BR}}$ for Brazil. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines for.

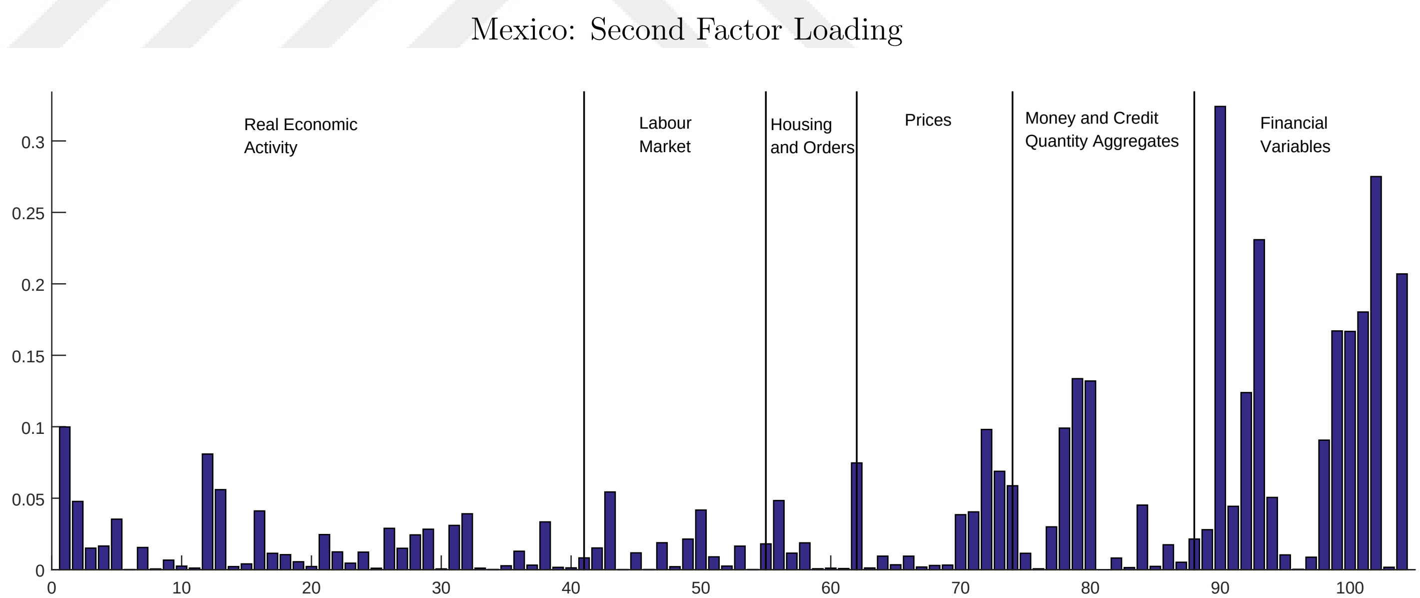


Figure 8: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{4,BR}}$ for Mexico. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

South Africa: Second Factor Loading

Figure 9: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{6},\mathbf{SA}}$ for South Africa. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

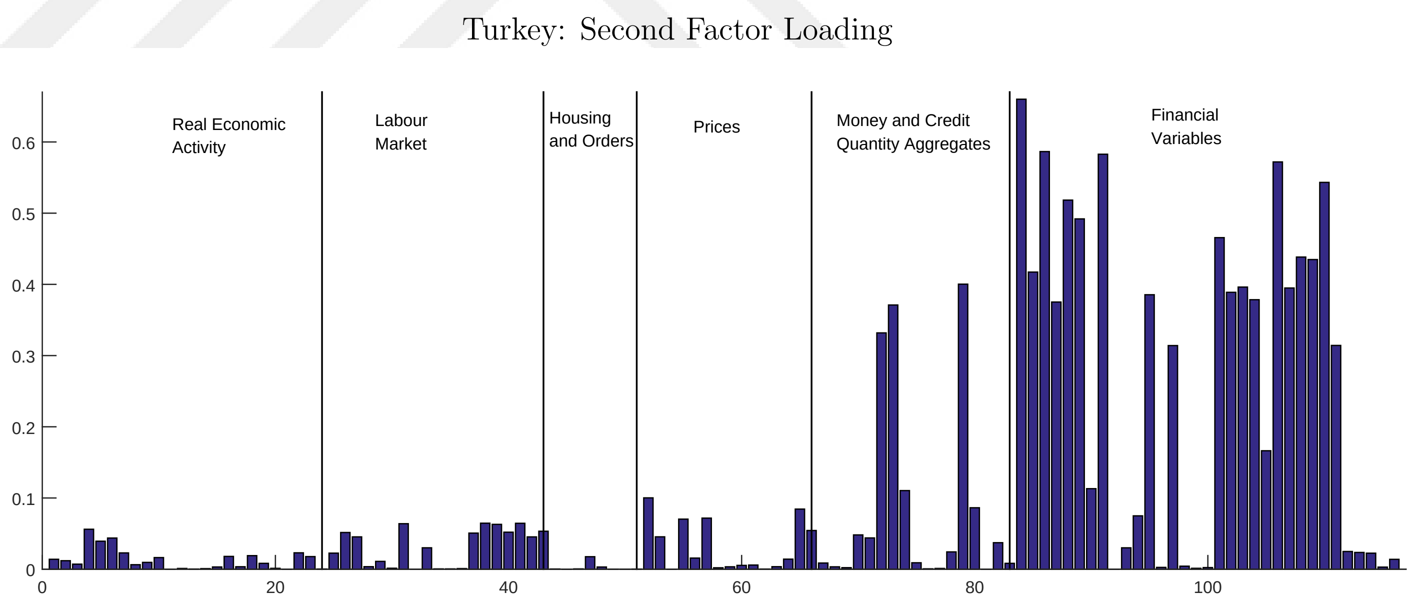


Figure 10: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{4,TR}}$ for Turkey. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

To have a better understanding of the estimated factor loadings, Figure 11 - 12 present the times series of the related factors (with economic interpretation) together with the main observed target variables. In Figure 11, I display the retrieved factor and contrast it with the various real economic activity variables. For instance, Panel A-B show extracted factors capture about 79% and 57% of the variation in industrial production variable for Brazil and Mexico respectively. Similarly, Panel C-D show that extracted factors capture about 65% of the variation in manufacturing production for South Africa and 78% of the variation in capacity utilization for Turkey. The main conclusion emerging from this figure is that, the respective factors for each country capture well the underlying low frequency dynamics in each of the observed series. This point us to the fact that EM local currency bond market risk premia bear direct relation to local macroeconomic fundamentals. Thus I call this *local real economic activity* factor. These findings are inline with the previous literature shows that emerging market asset returns are still substantially influenced by country specific factors. Figure 12 presents the retrieved factor for each country and contrasts them with the corresponding local currency exchange rate against the USD. Visual inspection shows that the retrieved factors capture well the high frequency dynamics of the target series. Average correlation of the retrieved factors with the currency variables is 71% in our sample. Impact of global financial shocks on aggregate fluctuations in emerging economies might be amplified as the exchange rate also responds to domestic fundamentals. Thus we might consider FX serving as a transmission mechanism of global financial conditions in local markets. [46] argues that an increase in global financial risk is an important channel through which the crisis is propagated to emerging economies. Recent literature also argues that emerging market bond risk premia are correlated with various global factors ([11] , [47]). Thus I call this the *global financial* factor.

To shed light on the underlying nature of these key identified factors, I also examine the cross-country correlation structure of the key factors and investigate the underlying sources of potential commonality. A number of surprising results emerge from this analysis. Table 3 shows while the global financial factor of each country highly correlated with each other, with correlation coefficients in the 39%-70% range. This result is important since it demonstrates how common dependence of this type could induce significant correlations among EM local currency debt return. In contrast, the local real economic activity factor of each country seems to move to the beat of its own drum, and is less correlated themselves, with correlation coefficients at or below -4%. This result may explain why recent literature find evidence of diversification benefit embedded in EM local currency debt markets and increasing share of LC debt in the portfolio of foreign investors. Positive correlation among global financial factor and negative correlations among local real economic activity are of key importance as well since the nature of these two estimated factors determines the characteristics of local currency debt returns in emerging sovereign debt markets and directly affects the ability of fund managers and other financial institutions to diversify the risk of global fixed income portfolios. Because portfolio theory implies that the correlation structure of local currency debt market across the emerging market scope should play a central role in determining global portfolio positions and influencing the flow of capital across countries. Thus, my findings point us to the fact that the portfolios of global investors in the EM local currency debt markets may be more diversified than is generally believed. Furthermore, the existence of a unique risk premium (via local real economic activity) in local currency debt returns validates the view that EM debt markets as a separate asset class. These results help explain why recent literature find some evidence of a separate risk premium embedded in EM local currency debt markets that make them less correlated with many asset classes (see [48]).

These findings are intuitive and consistent with the literature for two reasons. First, the asset-pricing theory states that all information should be included in prices and hence both the global factors and the country specific fundamentals should be reflected in the bond prices. I find that both the country-specific factor (local real economic activity) and the global financial factor are important determinants in LC excess return. Second, since the country specific fundamentals change slowly over time it should be the variation in the global financial factor that should be more reactive to global dynamics in driving local currency bond return.

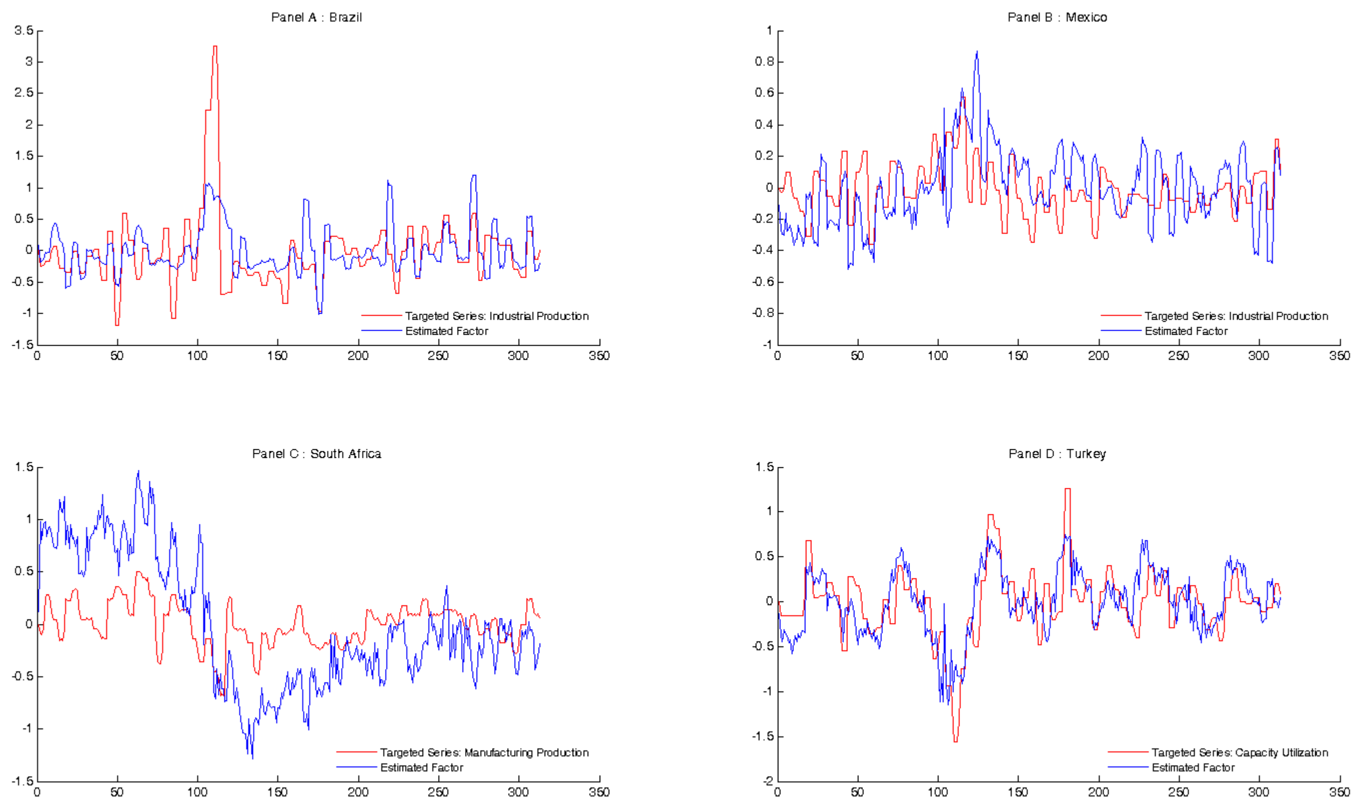


Figure 11: Time series of estimated factors (local real economic activity factors) against the targeted series. Estimated factors are plotted in blue lines and target series are plotted in red lines. Panel (A) plots Brazil target series (industrial production) versus estimated factor ($F_{\mathbf{4},\mathbf{BR}}$), Panel (B) plots Mexico target series (industrial production) versus estimated factor ($F_{\mathbf{5},\mathbf{MX}}$), Panel (C) plots South Africa target series (manufacturing production) versus estimated factor ($F_{\mathbf{1},\mathbf{SA}}$) and Panel (D) plots Turkey target series (capacity utilization) versus estimated factor ($F_{\mathbf{6},\mathbf{TR}}$).

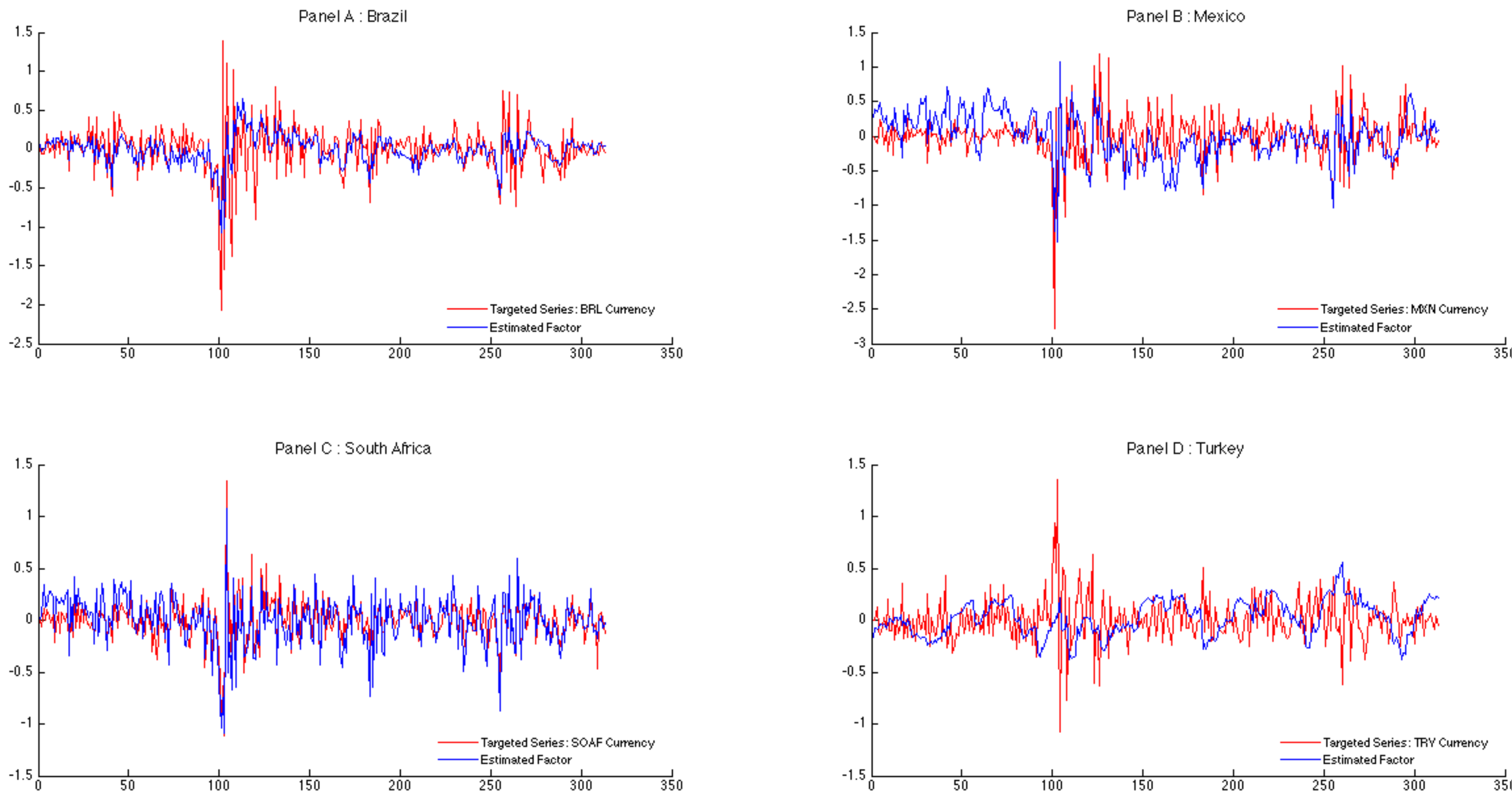


Figure 12: Time series of estimated factors (global financial risk factors) against the targeted series (local currency exchange rates against the USD). Estimated factors are plotted in blue lines and target series are plotted in red lines. Panel (A) plots Brazil target series (BRL/USD) versus estimated factor ($F_{\mathbf{2},\mathbf{BR}}$), Panel (B) plots Mexico target series (MXN/USD) versus estimated factor ($F_{\mathbf{4},\mathbf{MX}}$), Panel (C) plots South Africa target series (ZAR/USD) versus estimated factor ($F_{\mathbf{6},\mathbf{SA}}$) and Panel (D) plots Turkey target series (TRY/USD) versus estimated factor ($F_{\mathbf{4},\mathbf{TR}}$).

Table 3: Correlation Matrix of Target Factors: This table reports the pairwise correlation coefficients for local real economic activity factor and global financial risk factor for the indicated countries. Each pairwise correlation is computed using all available overlapping observations for the two sovereign.

Local Real Economic Activity Factor

	Brazil	**Mexico**	**S. Africa**	**Turkey**
Brazil	100%			
Mexico	-15%	100%		
S. Africa	-4%	-51%	100%	
Turkey	-46%	-14%	-28%	100%

Global Financial Risk Factor

	Brazil	**Mexico**	**S. Africa**	**Turkey**
Brazil	100%			
Mexico	43%	100%		
South Africa	60%	57%	100%	
Turkey	42%	39%	44%	100%

1.5.3 Forecasting Regressions

I now explore how much of the variation in LC excess return can be explained by extracted factors by running linear regressions as shown at Eq(7). From Table 5 to Table 8 presents results of forecasting regressions for two, three, four, and five year excess bond returns for each country. For each regression, I report the regression coefficients, t-statistics (based on the [49] heteroskedasticity-consistent estimate of the covariance matrix) and adjusted R^2 for each of the regressions.

Table 5 to 8 shows the results of predictive regressions for excess returns across the yield curve for each country. I find that linear combination of estimated factors with proposed specifications explain an economically large fraction of the variation in future returns. When we consider our results for the two year bonds, these factors are statistically significant predictors of bond returns and explain 24%, 31%, 44%, and 49% of the variation one year ahead in the two-year return for Mexico, Brazil, Turkey and South Africa respectively. More interestingly, as we move along the yield curve, we observe that forecasting power increases and these factors explain up to 34%, 43%, and 64% of next year's excess return on the three, four, and five year bonds respectively. The estimated factors have their strongest predictive power for five year bonds for South Africa. Notice also that the main predictor variables are factors based on real activity that are highly correlated with measures of industrial and manufacturing production, but factors based on global financial factors also contain information about future LC bond returns. The results reported in Table 5 to 8 point out to three important findings about EM local currency bond market. First of all, LC bond yields are time varying and are a quantitatively important source of fluctuations. Secondly, good forecasts of excess LC bond returns can be made with only a few estimated factors that summarize information from a large panel of local economic activity. Third, the factor for local real economic activity contributes most significantly to variations in excess bond returns for each country. To the best of our knowledge, my findings is one of the first to base the empirical analysis on data from a large and liquid market LC denominated debt of major emerging markets and show that macroeconomic variables can predict LC returns. Reported findings are also in line with the macroeconomic theory postulates that it is real variables relating to macroeconomic activity that should forecast bond returns and the empirical studies that find significant forecastable variation in the excess returns of U.S. government bonds (see [50] , [30] and Ludvigson and Ng(2009)).

I also form a single predictor function by fitting the values from a regression of average excess returns on the set of estimated factors for each country.[10] This single factor is empirically grounded in the results of [30] that take into account the presence of a common factor driving realized excess returns on US government bonds. My aim is to assess the forecasting performance of this single linear combination of the factors on excess bond returns at all maturities for all countries. For each country, the regressions are formed by using $\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)} = F_t$ as dependent variable and relevant factors as independent variables. I denote these single factors by $F_{BR}, F_{MX}, F_{SA}, F_{TR}$ respectively. Then we have the following single predictor form for each country:

[10] This single-factor model can be regarded as the single index models used by [51] and Robert Hodrick (1983) and [52].

Table 4: Single factor regressions for each country

Country	**Single Factor**	**Equation**
Brazil	$F_{BR,t}$	$\gamma_{1,BR}F_{2,BR,t} + \gamma_{2,BR}F_{3,BR,t} + \gamma_{3,BR}F_{4,BR,t}$ $+\gamma_{4,BR}F_{5,BR,t} + \gamma_{5,BR,t}F^2_{3,BR,t}$
Mexico	$F_{MX,t}$	$\gamma_{1,MX}F_{4,MX,t} + \gamma_{2,MX}F_{5,MX,t} + \gamma_{3,MX}F_{8,MX,t}$ $+\gamma_{4,MX}F_{11,MX,t} + \gamma_{5,MX}F_{15,MX,t}$
South Africa	$F_{SA,t}$	$\gamma_{1,SA}F_{1,SA,t} + \gamma_{2,SA}F_{2,SA,t} + \gamma_{3,SA}F_{4,SA,t}$ $+\gamma_{4,SA}F_{5,SA,t} + \gamma_{5,SA}F_{6,SA,t} + \gamma_{6,SA}F_{8,SA,t}$ $+\gamma_{7,SA}F^2_{1,SA,t} + \gamma_{8,SA}F^2_{5,SA,t}$
Turkey	$F_{TR,t}$	$\gamma_{1,TR}F_{1,TR,t} + \gamma_{2,TR}F_{4,TR,t} + \gamma_{3,TR}F_{6,TR,t}$ $+\gamma_{4,TR}F^2_{1,TR,t} + \gamma_{5,TR}F^2_{2,TR,t} + \gamma_{6,TR}F^2_{8,TR,t}$

Table 9 presents results for forecasting regressions for my general form of single predictor factors for each country. I report the regression coefficients, t-statistics (based on the [49] heteroskedasticity-consistent estimate of the covariance matrix) and adjusted $R2$ for each of the regressions. The results from these regressions reinforce those presented earlier andshow general form of single predictor factor explains between 34% to 57% of the variation in next years excess returns on LC bonds.

My results also have implication for the expectations theory (ET) of the term structure in emerging markets. ET states that variables in the information set F_t at time t should have no predictive power for excess bond returns. Conventional tests of ET for my predictive regression framework $rx_{t+1}^{(n)} = \beta F_t + \epsilon_t$ is to check the null hypothesis whether the parameter vector β is zero. Tables from 5 to 8 and Table 9 that show the predictive power of the estimated factors F_t is not just statistically significant but also economically important. Thus we can strongly reject that β is zero and claim that ET does not hold for term structure in Brazil, Mexico, South Africa and Turkey.

Table 5: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column. The dependent variable $rx_{t+1}^{(n)}$ is the excess log return on the n-year bond. F_t denote factors estimated by the method of principal components using a panel of data for Brazil. [53] corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

Brazil						
	$F_{2,BR,t}$	$F_{3,BR,t}$	$F_{4,BR,t}$	$F_{5,BR,t}$	$F_{3,BR,t}^2$	R-Square
rx(2)	0,1442	-0,2004	0,3715	-0,2420	0,0092	31%
	[**4,3684**]	**-**[**4,1927**]	[**7,7330**]	**-**[**4,2730**]	[1,3242]	
rx(3)	0,2477	-0,4144	0,7104	-0,6230	0,0142	34%
	[**4,0349**]	**-**[**4,6632**]	[**7,9530**]	**-**[**5,9155**]	[1,1015]	
rx(4)	0,2857	-0,2608	0,9993	-1,1306	0,0170	38%
	[**3,6951**]	**-**[**2,3300**]	[**8,8841**]	**-**[**8,5248**]	[1,0509]	
rx(5)	0,3591	-0,2623	0,8329	-0,6268	0,0671	43%
	[**4,0428**]	**-**[**2,0398**]	[**6,4453**]	**-**[**4,1140**]	[**3,6007**]	

Table 6: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column. The dependent variable $rx_{t+1}^{(n)}$ is the excess log return on the n-year bond. F_t denote factors estimated by the method of principal components using a panel of data for Mexico. [53] corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

	Mexico					
	$F_{4,MX,t}4$	$F_{5,MX,t}$	$F_{8,MX,t}$	$F_{11,,MX,t}$	$F_{15,MX,t}$	**R-Square**
rx(2)	0,1186	-0,1014	0,0822	-0,1325	-0,1532	24%
	[**5,6378**]	**-**[**4,5026**]	[**3,0253**]	**-**[**4,0336**]	**-**[**4,1385**]	
rx(3)	0,2967	-0,3154	0,2169	-0,2495	-0,3176	35%
	[**7,4005**]	**-**[**7,3425**]	[**4,1894**]	**-**[**3,9836**]	**-**[**4,5013**]	
rx(4)	0,3986	-0,5912	0,2104	-0,2603	-0,3450	35%
	[**7,6911**]	**-**[**10,6483**]	[**3,1436**]	**-**[**3,2149**]	**-**[**3,7825**]	
rx(5)	0,4050	-0,5895	0,3593	-0,3981	-0,3836	40%
	[**6,4804**]	**-**[**8,8037**]	[**4,4517**]	**-**[**4,0776**]	**-**[**3,4870**]	

Table 7: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column. The dependent variable $rx_{t+1}^{(n)}$ is the excess log return on the n-year bond. F_t denote factors estimated by the method of principal components using a panel of data for South Africa. [53] corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

South Africa

	$F_{1,SA,t}$	$F_{2,SA,t}$	$F_{4,SA,t}$	$F_{5,SA,t}$	$F_{6,SA,t}$	$F_{8,SA,t}$	$F^2_{1,SA,t}$	$F^2_{5,SA,t}$	**R-Square**
rx(2)	-0,1709	0,1821	0,2564	0,2373	-0,3152	0,2549	-0,0311	0,0294	49%
	-[**5,4620**]	[**5,6172**]	[**6,4443**]	[**5,6199**]	**-**[**6,9519**]	[**5,3035**]	**-**[**3,2314**]	[**1,9721**]	
rx(3)	-1,3580	0,0021	0,0210	-0,1350	-0,4063	0,8408	0,1003	0,0040	51%
	-[**15,8538**]	[0,0237]	[0,1933]	-[1,1680]	**-**[**3,2744**]	[**6,3922**]	[**3,8054**]	[0,0987]	
rx(4)	-0,2623	0,8157	0,7507	1,0036	-0,8305	-0,1856	-0,1411	0,0870	50%
	-[**5,3921**]	[**16,1852**]	[**12,1364**]	[**15,2918**]	**-**[**11,7853**]	**-**[**2,4842**]	**-**[**9,4211**]	[**3,7476**]	
rx(4)	-1,1423	0,4759	0,3908	0,4331	-0,5691	0,3403	0,0542	0,0722	64%
	-[**19,8397**]	[**7,9780**]	[**5,3386**]	[**5,5757**]	**-**[**6,8229**]	[**3,8488**]	[**3,0608**]	[**2,6300**]	

Table 8: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column. The dependent variable $rx_{t+1}^{(n)}$ is the excess log return on the n-year bond. F_t denote factors estimated by the method of principal components using a panel of data for Turkey. [53] corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

	Turkey						
	$F_{1,TR,t}$	$F_{4,TR,t}$	$F_{6,TR,t}$	$F^2_{1,TR,t}$	$F^2_{2,TR,t}$	$F^2_{8,TR,t}$	**R-Square**
rx(2)	-0,2078	-0,2756	0,6619	0,0460	0,0128	0,1516	44%
	-[3,8633]	**-[3,4693]**	**[6,8907]**	**[4,2753]**	**[2,1908]**	**[4,4121]**	
rx(3)	-0,3303	-0,4394	1,1399	0,1180	0,0285	0,3828	46%
	-[3,0882]	**-[2,7819]**	**[5,9679]**	**[5,5142]**	**[2,4496]**	**[5,6017]**	
rx(4)	0,4316	0,3718	0,2015	0,1219	0,0245	0,5461	51%
	[2,7507]	[1,6044]	[0,7190]	**[3,8809]**	[1,4397]	**[5,4474]**	
rx(5)	-1,0456	-0,9668	2,8686	0,1488	0,0520	0,2698	60%
	-[7,0526]	**-[4,4155]**	**[10,8338]**	**[5,0157]**	**[3,2282]**	**[2,8487]**	

Table 9: Single Factor Regressions: The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column. The dependent variable $F_{BR}, F_{MX}, F_{SA}, F_{TR}$ is the mean for the excess log return on two, three,four and five year maturities for indicated countries. [53] corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

Brazil	$F_{2,BR,t}$	$F_{3,BR,t}$	$F_{4,BR,t}$	$F_{5,BR,t}$	$F^2_{3,BR,t}$				**R-Squares**
	0,2592	-0,2845	0,7285	-0,6556	0,0269				34%
	[4,1894]	**-[3,1764]**	**[8,0938]**	**-[6,1777]**	**[2,0707]**				
Mexico	$F_{4,MX,t}$	$F_{5,,MX,t}$	$F_{8,MX,t}$	$F_{11,MX,t}$	$F_{15,MX,t}$				
	0,3047	-0,3994	0,2172	-0,2601	-0,2999				39%
	[7,6799]	**-[9,3954]**	**[4,2387]**	**-[4,1963]**	**-[4,2938]**				
South Africa	$F_{1,SA,t}$	$F_{2,SA,t}$	$F_{4,SA,t}$	$F_{5,SA,t}$	$F_{6,SA,t}$	$F_{8,SA,t}$	$F^2_{1,SA,t}$	$F^2_{5,SA,t}$	
	-0,7334	0,3690	0,3548	0,3847	-0,5303	0,3126	-0,0044	0,0482	57%
	-[17,8461]	**[8,6659]**	**[6,7891]**	**[6,9396]**	**-[8,9073]**	**[4,9536]**	[0,3542]	**[2,4571]**	
Turkey	$F_{1,TR,t}$	$F_{4,TR,t}$	$F_{6,TR,t}$	$F^2_{1,TR,t}$	$F^2_{2,TR,t}$	$F^2_{8,TR,t}$			
	-0,2880	-0,3275	1,2180	0,1087	0,0294	0,3376			47%
	-[2,8613]	**-[2,2030]**	**[6,7748]**	**[5,3947]**	**[2,6929]**	**[5,2487]**			

1.6 Conclusions

In this study, I depart from the existing empirical literature on EM sovereign foreign currency denominated debt and pay special attention to the local currency debt of four major emerging market countries; Brazil, Mexico, South Africa and Turkey over the period of June 2006 to March 2014. I use dynamic factor approach in which we exploit information from large economic and financial time series to assess: (i) the degree to which a small number of statistical factors, regardless of their nature, can be used to understand a broad set of economic indicators (ii) the degree to which the estimated factors, identified statistically, relate back to the set of macroeconomic variables, and (iii) the degree to which the estimated factors can predict local currency bond returns. I report a number of novel empirical results. First of all, I contribute to literature to show that strong predictable variation in the EM local currency excess bond returns that is associated with local macroeconomic activity. The adjusted R^2s for the forecasting regressions are fairly high, indicating that the estimated factors capture much of the variation in LC bond return. The lowest and highest values of the adjusted R^2s are 31% and 49%, respectively for two year maturity bonds. More interestingly forecasting power of factor augumented regressions increase as we move along the curve and these adjusted R^2s range from 40% to 64% percent for five year maturity bonds. Secondly, relating the estimated factors back to more economic variables, I find that the first factor seems to reflect local real economic activity and second factor seems to have more of global financial risk flavor. I provide the evidence that these two factors are both statistically and economically significant in explaining local currency excess return. These results highlight that it is important to include data-rich macroeconomic factors when forecasting EM local currency bond markets. This issue has fundamental implications for how the international fund manager should deal with local currency debt markets. I look forward to a variety of variations and extensions of my basic theme to: (i) construction of a real time composite leading index (ii) determination of the common driving factors and (iii) incorporation of the factors for developing an expert system for the assessment of monetary policies for major EM economies.

CHAPTER II

PREDICTABILITY OF EMERGING MARKET REAL ESTATE PRICES

2.1 Introduction

During the last decade, house prices in many emerging and advanced economies have moved synchronized with each other. Excessive credit expansions and overvalued exchange rates caused excrescent price booms in real estate markets of emerging countries until the onset of the global financial crisis in 2008. When the banks' massive losses based on the sub-prime mortgage bubble revealed, the problems in the housing market associated with the relaxed lending standards spread to the financial sector and led to a severe global financial crisis, which has taken its place in history as the Great Recession. House prices in many countries collapsed, and falling collateral values contracted borrowing capacity of households and firms in a procyclical manner. This recent crisis has drawn attention to the housing sector by emphasizing its importance as a determinant of macroeconomic activity and business cycle fluctuations. The current literature has primarily analyzed the link between real estate markets and output level both in emerging and developed economies by using some financial accelerator mechanisms. One of the main contributions of my study to the literature is to document the mutuality of top three factors predicting the real house price fluctuations in a sample of leading emerging economies including Brazil, Mexico, South Africa, and Turkey by using the dynamic factor analysis method. Moreover, with my empirical analysis I demonstrate that in terms of predictability, my model captures main shocks driving fluctuations better than any paper of which we are aware.

Due to the global decline in real interest rates, house prices in emerging markets underwent a substantial run-up until the peak recorded in early 2008, which is similar to that experienced in global financial markets. As shown in upper panel of Figure 13, the house price movements in emerging economies ally with the pattern displayed in the global landscape of housing from 2000 to 2015. After the global financial crisis that followed the collapse of Lehman Brothers, global house prices rapidly changed course and the house price growth rates both in emerging and global markets fell sharply due to the reversal of international financial conditions and investor expectations. Furthermore, real house price and GDP growth rates display a similar pattern for emerging markets and the United States between 2004 and 2015 as demonstrated in lower panel of Figure 13. The collapse of house price growth both in the United States and emerging markets is chased by a sharp decline in GDP growth rates. Since 2012, housing markets of the United States and emerging countries have started to rebound matching the pace of GDP for each country. Also, many emerging markets have seen striking increases in house prices over the past few years. In light of previous practices, these striking increases in house prices raise concerns about housing bubbles in the making, and the potential negative impact on financial stability and the overall economy.

There is a broad range of reasons to study dynamics of housing markets. First of all, housing makes up the largest component of wealth in many countries and homeownership rates are high across most of the OECD countries as displayed in Figure 14. For instance, in the United States, real estate account for roughly a third of the total assets held by the non-financial private sector and home ownership rate is about 65%. ([54]). A majority of households, especially in emerging markets, tend to hold wealth in the form of their homes rather than in financial assets. Also, in France, while less than a quarter of households own stocks, nearly 60% of households are homeowners. Secondly, since housing is the main asset and mortgage debt is the main liability held by households in many advanced countries, large movements in house prices can have serious macroeconomic implications by influencing households' capacity to borrow and allocate in residential investment. Next, consumption of housing and related expenses is the major driver of aggregate demand as it makes up a substantial fraction of GDP and household consumption. Finally, since the mortgage markets play a major role in the transmission of monetary policy, the weight of mortgage lending in banking system makes the value of housing assets critical for financial market stability. Therefore, we can assert that housing prices could account as an indicator of macroeconomic fluctuations and a well-functioning housing sector is critical for the overall health of the economy.

Housing market serves as a key barometer of the wider economy, capturing underlying pressures in the economic outlook. It is associated with changes in macroeconomic fundamentals not only because of its forward-looking nature that informs economic outlook by picking up factors such as income prospects and credit conditions, but also because it might directly cause changes in fundamentals of that outlook by affecting collateral values and thereby credit conditions. So, it is crucial to watch the developments in housing markets closely to be able to understand the potential concerns with regards to the economy.

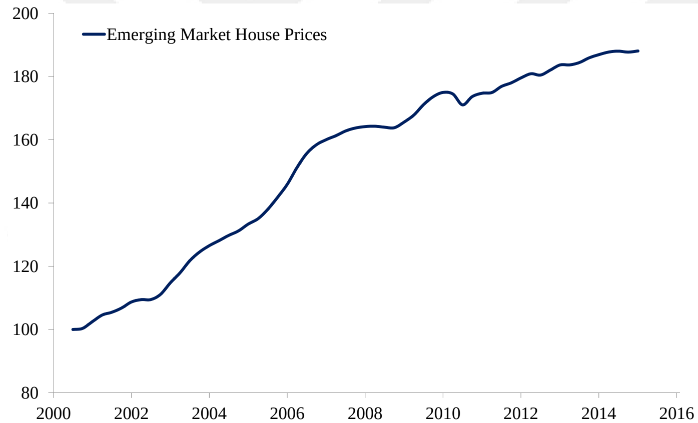
In this chapter, I investigate the characteristics of house price dynamics in four emerging countries including Mexico, Brazil, South Africa, and Turkey by using a new data set of real house prices. By linking the house prices to a set of determinants, I analyze the importance of different types of shocks in explaining movements of house prices. As a result, I find that a significant portion of the variation in real house price growth rates in sample countries could be explained by the same set of three factors, which contains financial variables, monetary and credit quantity aggregates, and real economic activity.

Economic theory suggests that house prices, rents, and incomes should move in tandem in the long-run equilibrium, emphasizing the significance of demand and supply shocks in housing markets. Therefore, the ratios of price to rent and price to income are most commonly used candidates to forecast house prices and to check the stability of macroeconomic fundamentals and housing prices. There is also a growing literature about identifying various shocks of different types that constitute a significant source of fluctuations in the housing market. However, forecasting house prices on a single or a few variables which may not cover all the space spanned by the structural shocks is quite inadequate and misleading. We need to combine all related sources of information as efficiently as possible with an overall view of the conditions in the market to proficiently forecast movements in housing prices.

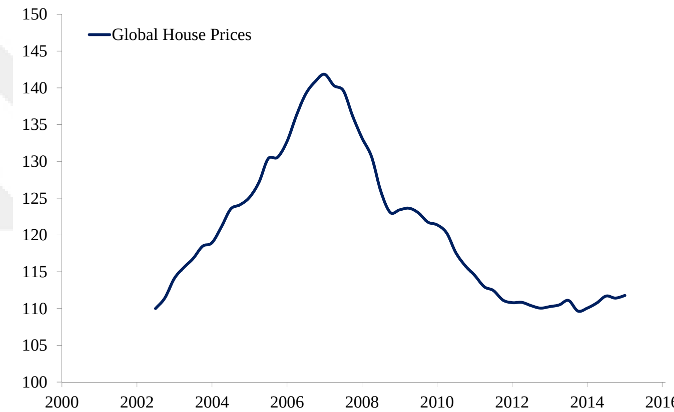
To overcome this problem, I use the factor model analysis which has received increased attention since the beginning of this past decade due to their suitability for analyzing large data sets. ([34] ; [55]; [29], [56]; [57]) Factor models help to identify the sources of fluctuations by inferring the number of shocks directly from the data using various tests and information criteria, rather than choosing sources on a priori grounds. Currently, factor models are used for forecasting ([28]; [58], constructing leading coincident indicators ([59]), structural analysis ([56]; [60]), and policy analysis. ([61]; [62]; [63])

I study a panel data of more than 100 economic time series for each country in my sample of four emerging economies including Mexico, Brazil, South Africa, and Turkey to examine the determinants of housing prices over the period of 2007: Q2 and 2015, which covers the bust of 2008 and the subsequent recovery. As a result of my empirical analysis, I find a number of novel conclusions about the nature of housing prices. First of all, a small set of three key factors, which consists of financial variables, money and credit quantity aggregates and real economic activity factors load heavily on variables, concluding that they have a substantial predictive power for real house price movements for all the emerging countries in my sample. Those three common factors cumulatively explain more than 50% of the total variation in real house price growth rates for each country. Besides the mutuality of top three factors predicting the real house price fluctuations in those emerging markets, the signs of the coefficients also suit the economic theory. Lastly, my study presents the top predictive results for the real house price fluctuations to the best of our knowledge.

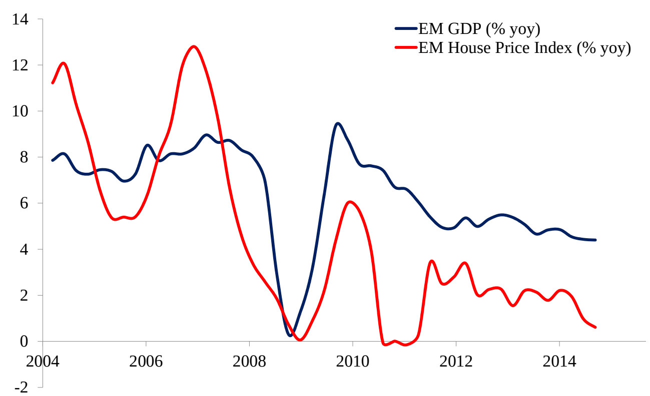
my findings also have important implications for policymakers and fund managers. As an essential part of the economy, the housing market has been the source of vulnerabilities and crises since housing is both an investment and consumption good and house purchases typically require debt financing. [64] documents that more than two-thirds of the almost 50 systemic banking crises in recent decades were preceded by boom-bust patterns in house prices. Hence, the decomposition of house price factor has important implications since it would help policymakers to implement market-specific diagnoses, and facilitate to find the right policy instruments that can ideally distinguish between underlying components of house price dynamics.



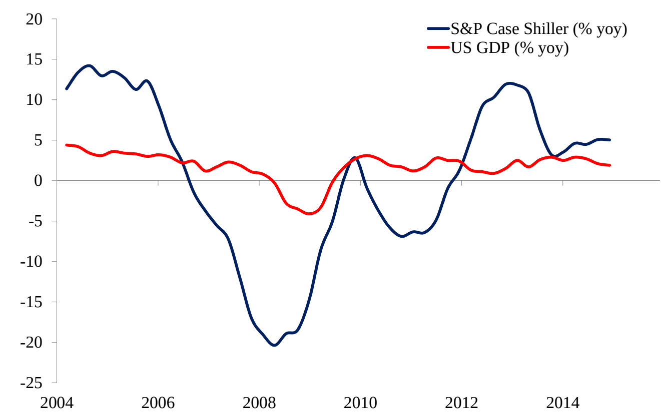
(a) Real House Prices: Emerging Markets



(b) Real House Prices: Developed Markets

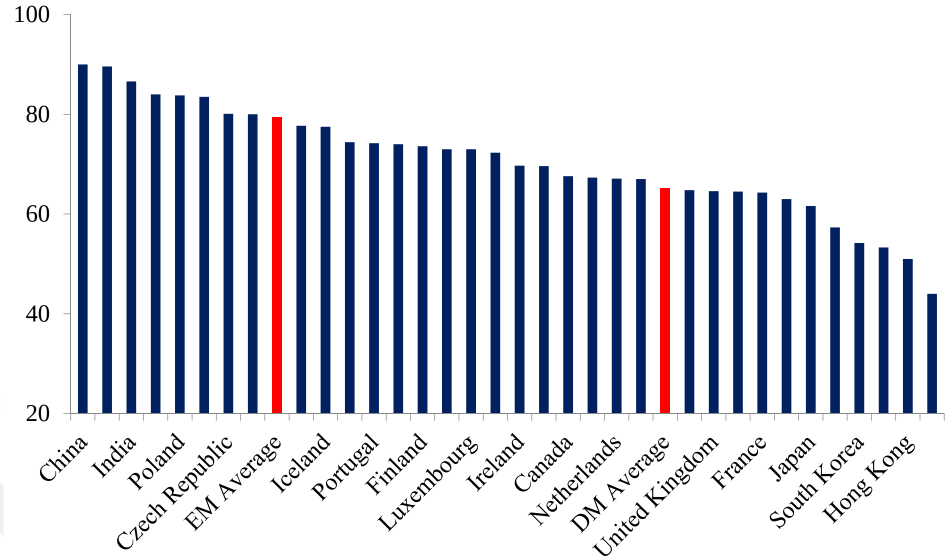


(c) Real GDP Growth in EM(yoy%) vs Real House Prices in EM (yoy%)

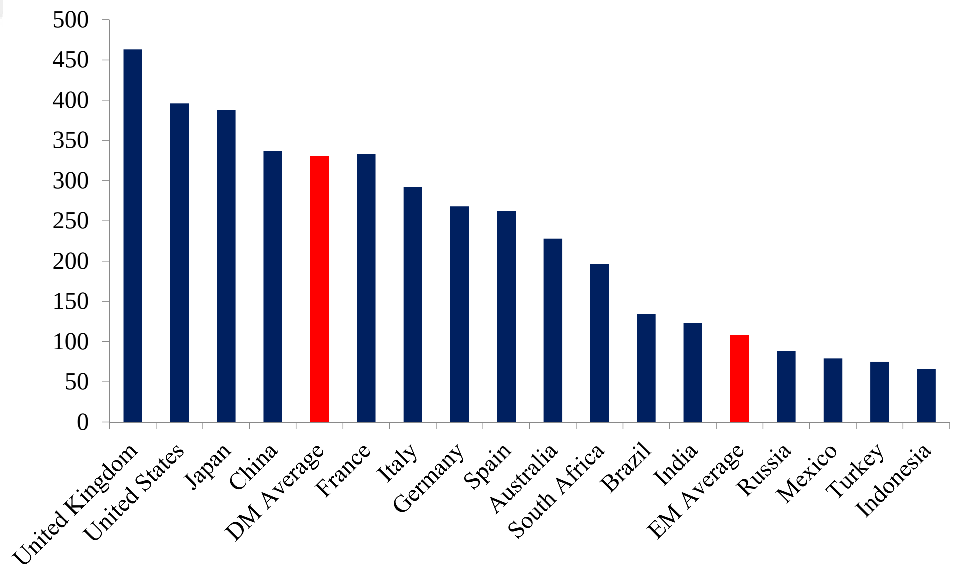


(d) US Real GDP Growth (yoy%) vs S&P Case Shiller House Price Index (yoy% - real series)

Figure 13: Inflation adjusted House Prices for both Emerging and Developed Markets are obtained from "OECD House Prices Database". For countries whose real house prices are not officially available, I use nominal time series and deflate them using consumer price indices published by national data resources.



(a) Home Ownership Rates



(b) Financial Assets to GDP

Figure 14: Panel(a) shows home ownership rates, as measured by the percentage of residential units that are occupied by their owners. These figures show that developed nations have accumulated more financial wealth while emerging economies hold a significant portion of their wealth in form of real estate.

2.2 Related Literature

There is a growing literature in house price fluctuation analyzing the importance of various shocks driving the global house market. In practice, it is difficult to know whether housing price movements are due to macroeconomic fundamentals that are tied to predictable components in the long-term demand and supply of the housing market and changes in credit conditions or irrational exuberance. Over past two decades, the literature splits into camps to understand the formation of house prices and housing price fluctuations. The first stream proposes to apply the asset-pricing framework ([1], [2], [3]) to housing values provide a framework by discounting a stream of rental prices and asserts that house prices, rents, and incomes should move aligned over the long run. Any deviation in house prices to rents ratio from its trend would encourage people to switch between buying and renting, eventually bringing the ratio back into its trend. Similarly, in the long run, if the house prices rise beyond people's affordability to buy them, the subsequent outcome is that house prices adjust to household income. Starting from [65], many studies show that housing returns exhibit positive autocorrelation and also document that they do not follow a random walk. Recent studies ([66] and [67]) also examine predictability in the housing market using the price-rent ratio. The fundamental problem with these studies is that proposed procedures treat many economic and financial variables as completely exogenous to the asset-pricing equilibrium. [68] review real estate booms and busts that concluded that investors work with simple heuristic models, instead of an extensive general equilibrium framework. Table 10 summarizes the key characteristics of reserach on house prices based on this literature stream.

As summarized in Table 11, the second literature stream has proposed a number of economic variables in order to provide a direct line linking house prices to economic fundamentals such as employment and income variables, ([69]), current account deficit ([70] and [71]), capital flows ([72] and [73]), collateral valuation ([74] and [75]) and real exchange rate ([76]). Relative to this strand of literature, I investigate the common latent factors that might have substantial predictive power for real house price fluctuations of four emerging countries by using a panel data of economic time series. Although I use an empirical strategy similar to [77], my findings differ by performing better in terms of predictability since I am looking at a different set of countries and a different set of economic time series. My research covered in this chapter, captures common factors that create main fluctuations in house markets of leading emerging markets, including Mexico, Brazil, South Africa, and Brazil, and present the best predictive results relative to a similar strand of the literature which we are aware of.

The third strand of literature attempts to explore the cyclical behavior of house prices and their relation to the macro economy for developed ([78], [79]) and emerging markets. ([80] and [81]) Similar to the previous literature, my research concentrates on house prices in four emerging economies and compares real house price movements systematically by using samples of quarterly data over a period covering the global financial crisis in 2008. In statistical terms, a significant percentage of variance in housing markets of those four emerging countries can be explained by three common factors calculated across a rolling window of the real house prices over the past quarters of 2007 and 2015.

Table 10: Literature Review in House Prices: Asset Pricing Framework

Authors	**Region**	**Data**	**Period**	**Methodology**	**Relationship**
Gallin (2008)	United States	Rent	1970-2001	Price/rent ratio	Positive
Case-Shiller (1988)	United States	House prices	1970-1986	Autocorrelation	Positive
Favilukis (2012)	United States	Rent	2000-2012	Stochastic General Equilibrium Model	Positive
Bracke (2011)	19 OECD Countries	House prices	1970-2010	Linear Probability Model	Positive
Campbell (2009)	United States	Rent	1975-2007	Dynamic Gordon Growth Model	Positive

Table 11: Literature Review in House Prices: Factor Models

Authors	Region	Data	Period	Methodology	Relationship
Andrews (2010)	OECD Countries	Interest rates, Disposable income, CPI and Employment	1980-2005	VECM	Income (+), Interest rates and unemployment (-)
Hilbers (2008)	European Union	HPI, Interest Rates, Demographics, Income and rents	1999-2003	User cost approach	User costs (-), Output (+)
Otrok et al. (2012)	18 Advanced Countries	GDP, House prices, Credit spreads,Uncertainty and Interest rates	1971-2011	Factor Augmented VAR	Rates (-), Uncertainty (-)
Tsatsaronis-Zhu (2004)	17 Advanced Countries	GDP, Interest rates, Spreads, CPI, Financing (Loans)	1970-2003	Structural Vector Autoregression (SVAR)	Inflation and Loans (+), Short term rates (-)
Goodhart-Hoffman (2008)	17 Advanced Countries	Money and Credit aggregates,Economic activity	1970-2006	Panel VAR	Money and Credit (+)
Adams-Füss (2010)	15 Advanced Countries	GDP, Interest rates and Real economic activity	1970-2006	Panel Regression	Economic Activity (+)
Iacovello (2011)	United States	Housing wealth, stock prices and consumption	1952-2010	General EquilibriumModel	Positive
Ciarlone (2012)	16 Emerging Countries	House prices, GDP, Stock prices,Employment,	1995-2011	Panel Regression	Wages (-), Rates (-),Stock prices (-)
Cesa-Bianchi et al. (2015)	57 Advanced and Emerging Countries	Capital flows, consumption, exchange rates and current account	1990-2012	Panel VAR model	Global liquidity shock (+)

The rest of the chapter is organized as follows. Section 2 describes my dataset used in the analysis and offers preliminary facts. Section 3 presents an overall framework for potential determinants and dynamics of the house prices by comparing house price characteristics of four emerging countries concerning their macroeconomic fundamentals and real economic variables. Section 4 outlines the main feature of the empirical strategy implemented to measure components of deflated house price movements and to quantify the contribution stemming from each of the factors. Section 5 reports the main results of the analysis obviously showing that top three factors predicting the real house price fluctuations in four emerging markets; Brazil, Mexico, South Africa, and Turkey, are common. Section 6 discusses the macro-prudential housing policies in Emerging Markets and Section 7 concludes. Additional information on the data and details of the analysis can be found in Appendix.

2.3 Data Description

My empirical investigation concentrates on house prices for four major emerging markets—namely, Brazil, Mexico, South Africa, and Turkey. My choice of emerging market countries is mainly constrained by the lack of sufficient numbers of housing data and available economic series and the data I use in this study are imperfect. The major obstacle is that neither the house prices nor the rent data accurately price changes and data frequency can be inconsistent. So I refer to the latest results from the literature on the measurement of house prices and rent to adjust the published data and also to select countries to analyze. ([66])

Housing Price Data: My empirical investigation concentrates on house prices for four major emerging markets—namely, Brazil, Mexico, South Africa, and Turkey. My choice of emerging market countries is mainly constrained by the lack of sufficient numbers of housing data and available economic series and the data I use in this study are imperfect. The major obstacle is that neither the house prices nor the rent data accurately price changes and data frequency can be inconsistent. So I refer to the latest results from the literature on the measurement of house prices and rent to adjust the published data and also to select countries to analyze. ([66])

Rent Data: My primary source for rent data is the index for tenant's rent from the Consumer Price Index (CPI). Similar to [66], I use CPI since it is a measure of the rent that owners implicitly pay to themselves. Using rent related sub-indices of CPI is reasonable as it provides a relevant measure of the rental costs in an economy. In other words, it is like "dividends" for owners. CPI based rent series has significant advantages for my research. First, it is available for longer time series and time series for rents are not available for many of the emerging market economies.

Economic and Financial Series: I estimate factors from a balanced panel of 124, 111, 108, and 118 economic series for Brazil, Mexico, South Africa, and Turkey, respectively. This is the same dataset that I use to identify empirical linkages between macroeconomic variables and EM Local currency bond returns. The same two-step process is applied to adjust the data before analysis. First , these series are transformed into stationary series and then, these series are normalized by subtracting the mean and dividing by the standard deviation to have a zero sample mean and unit variance. This standardization is necessary to avoid overweighting of the series with large variance.

I use the same data classification for the analysis of house prices dynamics. Data is divided into the same sub categoreies; **Real Economics Activity Variables:** Industrial production, retail sales, international trade, and car sales. **Housing Variables:** House price indices, building permits and real estate units sold. **Labor Market Variables:** Employment and unemployment. **Prices:** Consumer prices, Producer prices and commodity prices. **Money Credit Quantity Aggregates:** Monetary base, money supply (M1-M4), and deposits (time, demand, and foreign exchange). Financial Variables: Interest Rates, Exchange Rates, and Stock price indices.

2.4 Dynamics of the House Prices:

Table 12 reports key statistics for the nominal house prices for all countries. I report average, median, standard deviation and autocorrelation. House price is high and very volatile in emerging markets. This observation is consistent with the faster growth of output and consumption in emerging economies. Similar to findings of [65], Table 12 also shows that housing returns exhibit positive autocorrelation and document that housing prices do not follow a random walk. Figure 15 provides a visual characterization of the house price cycle in selected economies. As we can see from Figure 15, countries have seen striking increases in nominal house prices over the last decade. From 2007 to the present, nominal house prices have risen by 321% in Brazil, 141% in Mexico, 132% in South Africa and 165% in Turkey. What is more impressive is that house prices in these countries did not experience substantial declines during the Lehman Collapse and European Sovereign Crises.

Table 12: Nominal House Prices: Key Statistics and Correlation Matrix

Panel A. Key Statistics for Nominal House Prices

	Brazil	Mexico	South Africa	Turkey
Average Monthly Return	1.23%	0.36%	0.30%	0.53%
Median	1.28%	0.32%	0.23%	0.75%
Standard Deviation	2.06%	1.64%	1.90%	3.53%
Autocorrelation	100%	60%	91%	83%

Panel B. Pairwise Correlations between Nominal House Prices

	Brazil	Mexico	South Africa	Turkey
Brazil	100%			
Mexico	78%	100%		
South Africa	88%	97%	100%	
Turkey	99%	84%	92%	100%

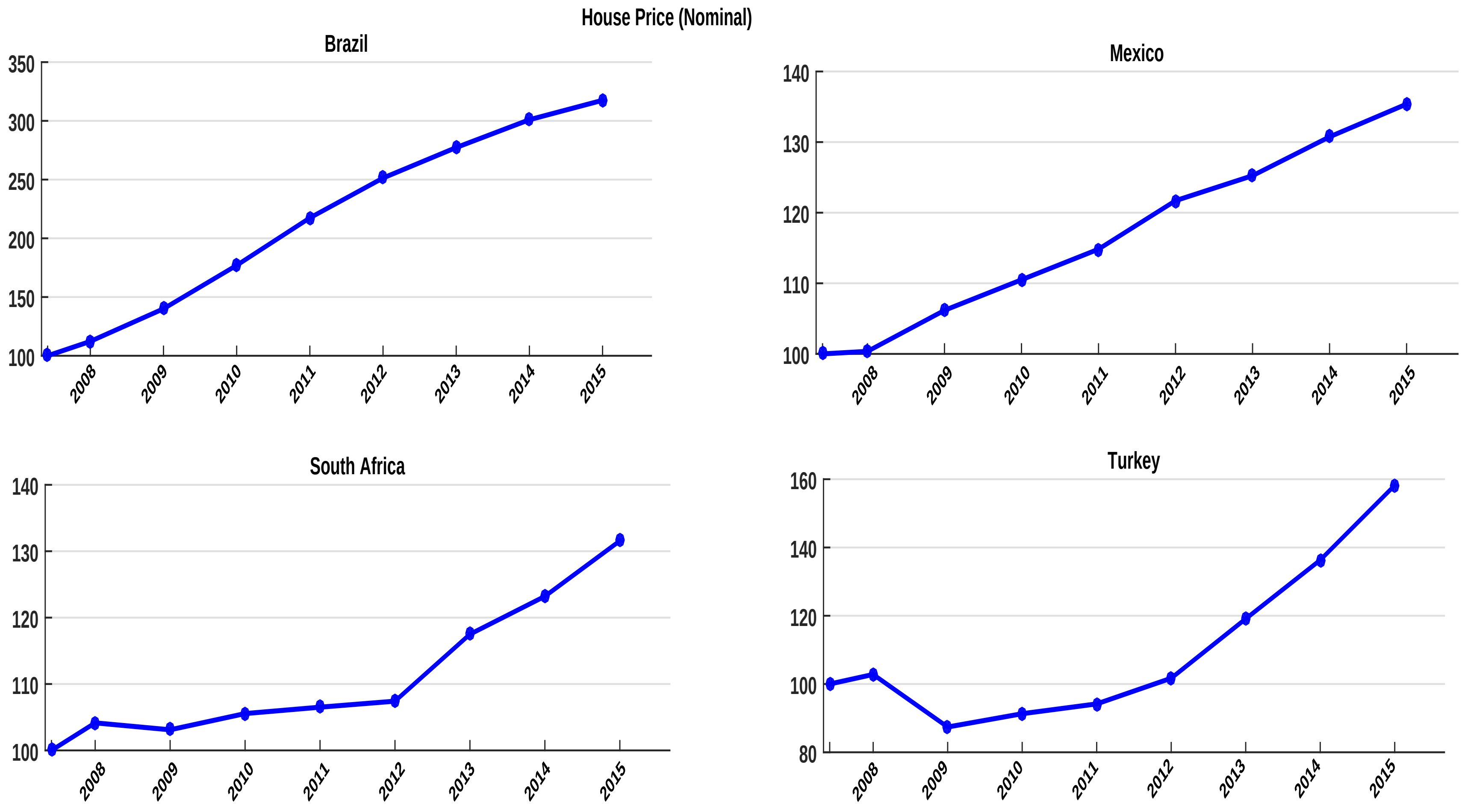


Figure 15: Nominal House Prices. I use national data resources as given in OECD House Prices Database. One exception is Turkey where I use data published by a private source (REIDIN) which has a longer history. All time series are normalized at 100 by June 2006.

Consumption of housing and related expenses are one of the major drivers of aggregate demand in all these economies thus housing prices are essential to account for macroeconomic fluctuations. Reported significant increases in home prices may raise concerns about the housing bubble and the potential negative impact on financial stability. However, an entirely different picture emerges when I use the real (deflated[1]) house price growth, which is given by log nominal house price growth minus log inflation. As seen from Figure 16, over the period 2007–2015, real house prices declined 12% and 18% in Turkey and South respectively and grew only 1% in Mexico and 208% for Brazil. Fluctuations that we observe in real housing prices are relevant not only to macroeconomic volatility and financial market stability but also consumption, wealth accumulation, labor mobility.

Lower panels of Table 13 and 12 report the pairwise correlations in nominal and real terms as measure of synchronization across countries. Cross country correlations in nominal house prices are very high, ranging from 78% to 97%. Reported negative cross country correlations in real house prices show that opposite holds for nominal house prices. This statistics shows that while nominal house price is synchronized across countries however real house price display little commonality over this period.

[1]To deflate nominal house price series into real terms, I am using appropriate price index for each country. The nominal data series is simply the data measured in local currencies and gathered by either a public source (i.e. Central Banks) or a private survey. Among the more prominent price indices, I use the Consumer Price Index (CPI) as the deflator.

Table 13: Real House Prices: Key Statistics and Correlation Matrix

Panel A. Key Statistics for Real House Prices

	Brazil	Mexico	South Africa	Turkey
Average Monthly Return	0.77%	0.01%	-0.20%	-0.11%
Median	0.81%	-0.06%	-0.16%	0.06%
Standard Deviation	2.05%	2.71%	2.42%	4.57%
Autocorrelation	89%	47%	72%	56%

Panel B. Pairwise Correlations between Real House Prices

	Brazil	Mexico	South Africa	Turkey
Brazil	100%			
Mexico	-41%	100%		
South Africa	-88%	71%	100%	
Turkey	-9%	-1%	5%	100%

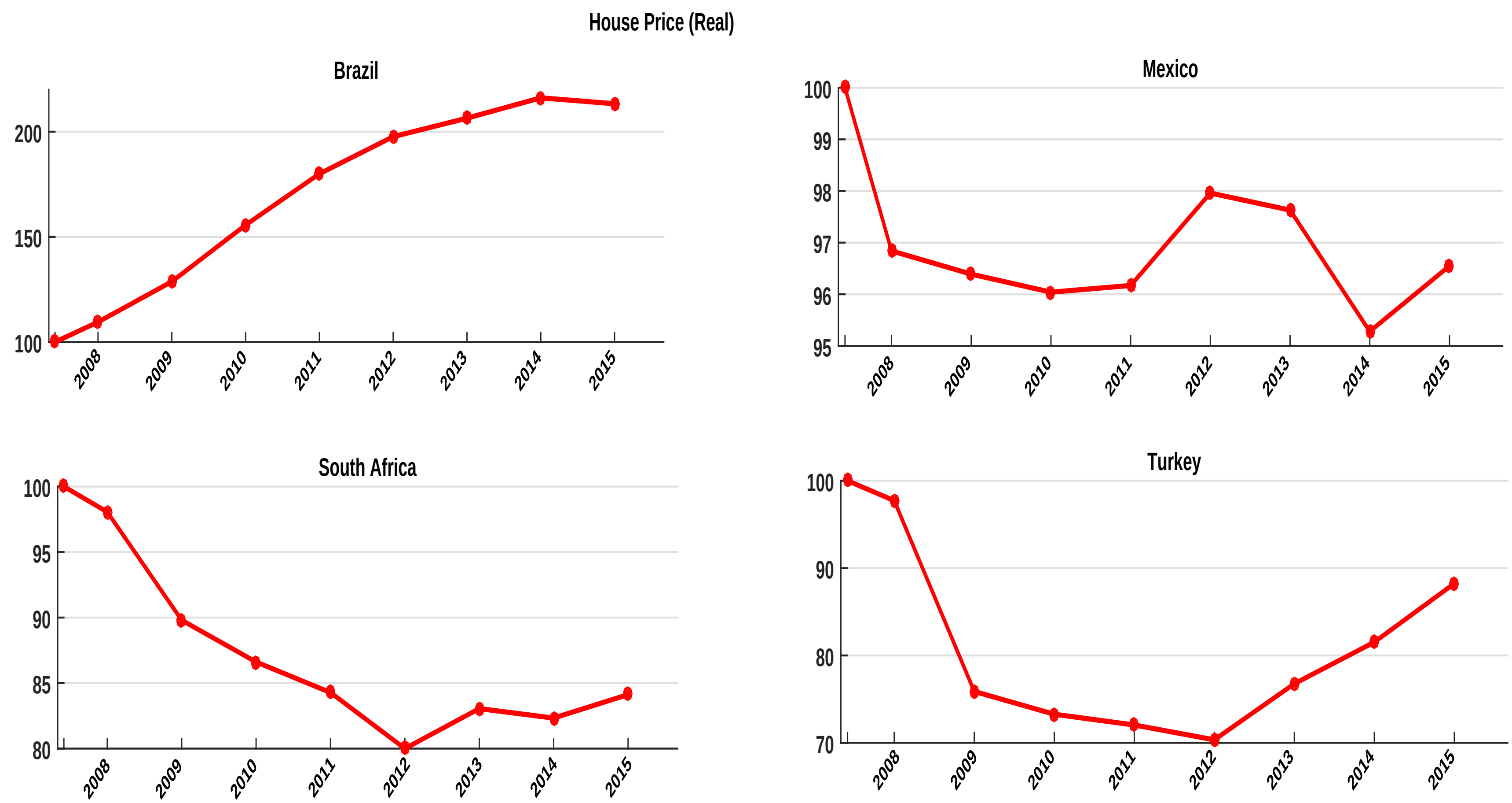


Figure 16: Real House Prices. I use national data resources as given in OECD House Prices Database. One exception is Turkey where I use data published by a private source (REIDIN) which has a longer history. I deflate nominal house prices series with national consumer price indices to calculate real price series. All time series are normalized at 100 by June 2006.

2.5 Empirical Methodology

In this section, I first compare the widely - known dynamic factor model and collapsed factor model. Then I represent the specifications of model structure, factor estimation and forecasting results.

2.5.1 Dynamic versus collapsed factor models

Many different methodologies exist for the purpose of forecasting the house prices in emerging markets. These can include simple bridge models or more sophisticated dynamic factor models. Over the last decade, use of the dynamic factor models become widespread among practitioners and econometricians due to the possibility of exploiting more data in the analysis and their good forecast performance. These new generation factor models differ from the strict factor models in which idiosyncratic errors are uncorrelated at least three important ways: (i) the number of observations is large (in both the cross-section (N) and the time (T) dimensions) which opens the horizon for consistent estimation of the estimated factors. (ii) the idiosyncratic errors can be serially and cross-sectionally correlated which makes this framework suited for a wider range of economic applications. (iii) Despite the increasing attention for dynamic factor models, choosing the appropriate dynamic factor model specification is still a topic of ongoing debate.

As opposed to daily financial series like bond and equity prices, we have monthly series of house prices and these monthly series may contain so-called jagged edges at the beginning and the end of the sample. Thus, it is crucial to pre-treat the jagged edges in the house price dataset before extracting the factors. I concentrate on the recently proposed collapsed dynamic factor (CDF) model of [82] due to its superior performance efficient handling of monthly series with different publication delays, and different starting dates. CDF model is a sibling of the canonical factor model of [57] and is based on the idea of using principal components to summarize the information in a large set of economic and financial time series. Nonetheless, in contrast to [57], CDF model estimates the target and the principal component variables concurrently in a low-dimensional multivariate unobserved component time series model to take account of panels with mixed frequencies and missing observations. This allows me to balance the data set and ease the estimation procedure for factor extraction.

[83] describe the econometric groundwork of the collapsed dynamic factor model in two steps. In the first step, the principal components are computed, and its dynamic properties are estimated using a vector autoregressive model. In the second step, the factor estimates and forecasts are obtained from the Kalman filter and smoother. The model of [82] differs from [83] in the following regards: Firstly, CDF follows low-dimensional unobserved components model for the target variable and a set of principal components from which the dynamic factors are extracted. The unidentified parameters in this parsimonious model are jointly estimated by using maximum likelihood for which the log likelihood function is evaluated using the Kalman filter and smoother. This structure clearly gathers all cross-sectional and dynamic time series information in an optimal way. Also, the idiosyncratic part of the target vector series is modeled explicitly and estimated jointly with the dynamic factors. This approach alleviates the general problem that the estimated factors extracted from a large macroeconomic panel are disregarding the information from the forecasting target.

Recent studies indicate there might be a promising extension regarding forecast accuracy by including autoregressive terms of the target variable in the model specification. In my case, this may include adding one or more lags of the targeted variable, House Prices ([77]), in the forecast equation.[2] The equations are cast in state space, enabling efficient handling of the jagged edges in the data and allow easy forecasting

[2] [84] ,and [85] find that forecast accuracy of simple bridgeequations can be significantly improved by the inclusion of an autoregressive term

via the Kalman filter and smoother.The recently proposed collapsed dynamic factor model shows the highest forecast accuracy for the macroeconomic panels with shorter time and cross-sectional dimension which is of great importance for emerging markets where some macroeconomic time series are less extensive.

2.5.2 Model Specification and Extraction of Factors

I follow [82] and use Collapsed Factor Model in my house price forecast experiment. I first consider a panel of observable economic variables $X_{i,t}$, where i denotes the cross-section unit $i = 1; ...; N$, and t refers to the time index $t = 1; ...; T$. Then I transform these series into stationary variables with zero mean and unit variance and label this data set as $x_{i,t}$. I use deflated house price growth as target variable which I obtain by log nominal house price growth minus log inflation and define the log real house price growth rate as $y_t = 100 \times (\ln(P_t/P_{t-1}) - \ln(CPI_t/CPI_{t-1}))$ where P_t and CPI_t are nominal house price and consumer price index at time t respectively. I specify the dynamic factor model as

$$\begin{aligned} y_t &= \mu_t + \Theta_{Fy} F_t + \epsilon_t \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \\ x_t &= \Lambda F_t + e_t \quad e_t \sim N(0, \Sigma_e) \end{aligned} \tag{6}$$

F_t in equation (1) is the set of latent factors that reflect most of the co-movement in the economy, Θ_{Fy} and Λ are factor loading matrix on y_t and x_t respectively. Error terms e_t and ϵ_t are assumed to be serially independent and independent of each other over time. Since estimation of parameters is not computationally feasible with large number of variables, I apply principal component analysis as a as a standard tool to reduce dimensionality and collapse my dynamic factor model with matrix A as $\hat{F}_{PC,t} = Ax_t$. The matrix A contains eigenvectors corresponding to one of the r largest ordered eigenvalues of xx'. Constriction of matrix A as $A\Lambda = I_r$ satisfies the condition of orthogonal factor extraction where $F_t = \hat{F}_{PC,t} + error$. $\hat{F}_{PC,t}$ is r dimension vector of principal components as $r << N$ and error is the discarded information for common component. Considering efficiency problem of principal component estimates with the exception of exact factor model accompanied by homoscedastic idiosyncratic components, I convert the model to linear state space form and perform Kalman Filter and smoother so as to utilize all the information on the present and past observations. The joint model enables us to extract out all the useful information for target variable which can be shown as

$$\begin{pmatrix} y_t \\ \hat{F}_{PC,t} \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} y_t^* \\ F_t \\ F_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \tilde{e}_t \end{pmatrix} \tag{7}$$

$$\begin{bmatrix} 1 & -\Theta_{Fy} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_{t+1}^* \\ F_{t+1} \\ F_t \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi_1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} y_t^* \\ F_t \\ F_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t \\ \xi_t \\ 0 \end{pmatrix} \tag{8}$$

where $\tilde{e}_t = A(e_t - [A' - \Lambda]F_t)$ is the macroeconomic stochastic shocks to the common factors and $Var(\tilde{e}_t)$ is an unknown matrix and needs to be estimated. Also the error terms ϵ_t and $\tilde{e}_t$ are assumed to be serially independent and independent of each other over time. I specify the latent variables F_t as $AR(p_\psi)$ and $VAR(p_F)$ respectively and set the state space representation of the collapsed dynamic factor model for $p_\psi = p_F = 1$ in order to capture the lead and lag relations among series along business cycles.

I use a methodology developed by [34] to estimate the optimal number of common factors by proposing some exclusion criteria under the assumption of large cross-section, N , and large time dimension, T. Because of different choices of the penalty function, Bai and Ng propose three criteria to determine the correct number of common factors and I use one of these criteria which gives me the least amount for the correct number of factors. From these estimated factors $\hat{f}_{PC,t}$, where $\hat{F}_{PC,t} \subset \hat{f}_{PC,t}$, we set up my model which we have shown in equation (2)-(3). I choose the preferred set of factors $\hat{F}_{PC,t}$ by regressing each estimated factors on target variable and minimizing the Bayesian information criterion (BIC) where I aim to pick factors with most predictive power for our forecasting process.

My analysis is based on more than a hundred indicators for each EM country where several indicators are observed in different time periods or missing and the model is unable to include jagged edge in x_t but only contains principal components $\hat{F}_{PC,t}$ in state space setup. I apply a stationary $AR(1)$ model for each indicator in x_t separately in order to handle the missing values at the beginning and at the end of x_t separately because principal component analysis is impractical with an incomplete set of observed data. I estimate the parameters of $AR(1)$ specification by using Maximum Likelihood Estimation in state space framework and use Kalman Filter and Smoother. Table 14 shows factor specifications for each country.

There are different approaches to the construction of forecasting process. One of the most popular methods is developed by [28]. In this method forecasting is based on autoregressive model where coefficients are estimated by least square method and re-estimated for each new observation in an iterative process. I prefer to stick to state space model for my forecasting procedure. I place selected factor to my state space model and run Kalman filter for out-of-sample observations. In this way, we exploit prediction step of Kalman filter and ignore the correction step.

Table 14: Factor specifications for each country

Country	Selected Factors (BIC Criterion)
Brazil	$F_{2,BR}$,$F_{4,BR}$,$F_{6,BR}$
Mexico	$F_{1,MX}$,$F_{3,MX}$,$F_{4,MX}$,$F_{12,MX}$,
South Africa	$F_{2,SA}$,$F_{4,SA}$,$F_{10,SA}$,$F_{12,SA}$,
Turkey	$F_{1,TR}$,$F_{2,TR}$,$F_{4,TR}$,

2.6 *Empirical Results*

2.6.1 Empirical Analysiss

In this section, I present my empirical results starting with the extraction of latent common factors from the panel that lead to a majority of the variation in the economic series and discuss the economic interpretation of these key factors. I provide results of my predictive regressions for real house price growth on common factors and the price-rent ratio as variables. Finally, I conduct a forecasting experiment on deflated house price growth rates based on a recursive scheme using all available data at the time of the forecast and report the outcomes of the out-of-sample analysis.

2.6.2 Extraction of the Common Factors

Using the methodology of Information Criteria (IC) developed by [34], I find that the factor structure of the panel is well described by a few number of factors, with zero mean, unit variance and which are mutually orthogonal to each other by construction.

Figure 17 demonstrates that in each of these four major emerging markets, a relatively small number of factors can explain more than 50% of the total variation in economic series from the panel over the entire span of time. The criteria indicate that the factor structure is well described by twelve common factors for Mexico, South Africa, and Turkey and six common factors for Brazil. Also, the first factor for each country explains the largest fraction of the total variation in the panel of data and the incremental power of each additional factor declines quite sharply. For example, the first factor for Mexico explains 18% of the variation in the panel, with the second factor explaining an incremental 11%, cumulating up to about 29%, and so on. If we check the variance contributions of first five factors selected by the IC, we see that they account for almost 50% of the variation in the economic time series. From these estimated common factors ft, I form a range of possible specifications for the forecasting regressions of real house price growth rates by using the Bayesian

Information Criteria (BIC). While the use of dynamic factor analysis with IC allows us to have a much larger set of predictor factors, the BIC provides an efficient way of choosing among summary factors by indicating whether these variables have important additional forecasting power for real house price growth rates. I choose the preferred set of factors Ft which minimizes the Bayesian Information Criterion (BIC).

Figure 18 , reports the results of my selection of a subset of common factors that are most informative about future deflated house price growth for each country out of the small set of factors extracted from the data. We see from the Figure 18 that three factors for Brazil and Turkey and four factors for Mexico and South Africa are optimal choice sets in terms of minimizing Bayesian Information Criteria (BIC). These factor representations account for about 60%, 55%, 52%, and 62% of the variation in the real house price growth rate series of Brazil, Mexico, South Africa, and Turkey, respectively. In my empirical analysis, I mainly focus on the first three common factors, which cumulatively explain more than 50% of the total variation in real house price growth rates for each country. Hence, we consider $F_{BR,2}$, $F_{BR,4}$, $F_{BR,6}$ for Brazil; $F_{MX,1}$, $F_{MX,3}$,$F_{MX,4}$ for Mexico; $F_{SA,2}$, $F_{SA,4}$, $F_{SA,10}$ for South Africa, and $F_{TR,1}$, $F_{TR,2}$, $F_{TR,4}$ for Turkey.

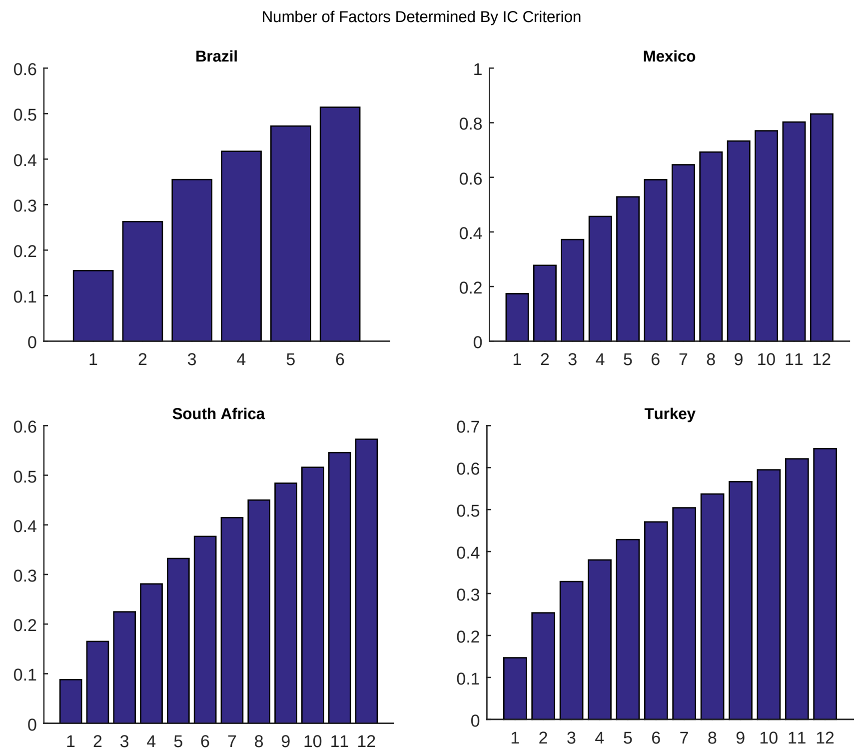


Figure 17: This figure shows overall variance contributions of factors selected by the IC for the indicated countries.

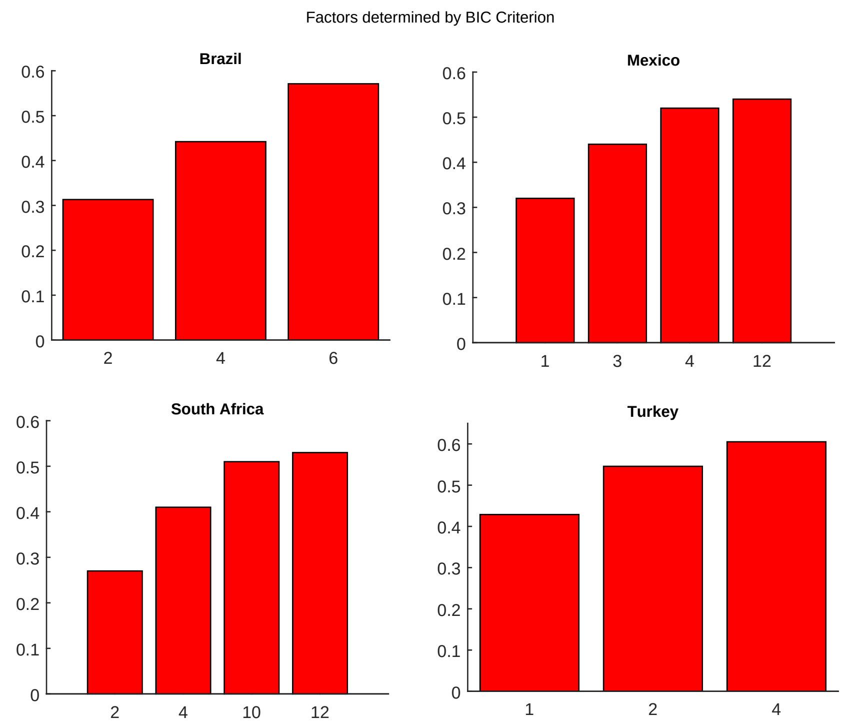


Figure 18: This figure shows the R^2 contributions of linear factors selected by the BIC for the indicated countries.

2.6.3 In-sample Results

After finding key factors that contain substantial information about the future movements of real house prices, I investigate if the estimated factors represent the common business cycle of the original variables. In general, the interpretation of the factors as representing specific macroeconomic or financial series is inappropriate since the construction of each factor is affected to some degree by all the variables in the large panel of data. Thus, none of the common factors correspond exactly to a precise economic concept like production, unemployment or interest rates. Furthermore, since the estimation of the factors allows identifying common factors only up to a matrix of constants, what matters is the dynamics of the factors over time rather than their specific values. With this important caveat, I follow the methodology of [28] and [86] in characterizing the factors and relating them to each of the variables in my panel dataset. I accomplish this by estimating the marginal R^2 statistics of univariate regressions of each economic time series variable included in the data set on each of the factors separately. It is critical to obtain an economically meaningful identification of the estimated factors to get an intuition of what information the factors might summarize.

Figures 19 to 30 show the R^2 statistic as bar charts from regressions of each of the individual series in my data set onto each estimated factor, one at a time for each country. Figure 19 to 22 suggest that $F_{BR,2}$, $F_{MX,1}$, $F_{SA,2}$, $F_{TR,1}$ which explain the largest fraction of the total variation in real house price growth rates load heavily on measures of financial variables for all of the four emerging countries. These factors explain up to 75%, 70%, 70%, and 65% of the total variation in the financial variables of Brazil, Mexico, South Africa, and Turkey, respectively. Second-factor components which are $F_{BR,4}$, $F_{MX,3}$, $F_{SA,4}$, and $F_{TR,2}$ for Brazil, Mexico, South Africa and Turkey, respectively load heavily on money and credit quantity aggregates. These factors explain up to 42%, 32%, 35%, and 40% of the total variation for the monetary and credit quantity aggregates for Brazil, Mexico, South Africa, and Turkey, respectively. Third-factor components for these emerging markets, which are listed as $F_{BR,6}$, $F_{MX,4}$, $F_{SA,10}$, $F_{TR,4}$ loads heavily on real economic activity. These factors explain up to 50%, 60%, 65%, and 60% of the total variation in the financial variables for Brazil, Mexico, South Africa, and Turkey, respectively. Hence, my findings point out that a small set of three key factors, which consists of financial variables, money and credit quantity aggregates, and real economic activity factors load heavily on macroeconomic variables and might have substantial predictive power for real house price growth rates for Brazil, Mexico, South Africa, and Turkey. One noteworthy feature of my study is its ability to capture common factors that create main fluctuations in house markets of leading emerging economies and present the top predictive results with higher levels of R^2.

Now, I investigate the predictive power of common factors for real house price growth rates based on the following predictive regression, which allows me to study the unconditional predictive power of the specified factors:

$$y_t = \mu_t + \theta_{pr,t} + \epsilon_t \tag{9}$$

Finally, I report results of in-sample tests for each country that demonstrate the predictive ability of each factor and price-rent ratio for real house price growth rates in Table 15 .

First, let's consider the top panel of Table 15, which shows the results of predictive regressions for real house price movements in Brazil. Since the price-rent ratio is the most popular housing market predictor and has received widespread attention in the literature, I accept price-rent ratio as a benchmark. Row (b) reports the results from a forecast specification including only the price-rent ratio as the predictor variable. Given the regression results, this variable is statistically significant and explains 24% of the variation in real house price growth rates. By comparison, row (a) presents that three-factor model is a stronger predictor of real house price fluctuations with statistically significant coefficients and an R^2 of 57%. When we add the price-rent ratio into the three-factor model as in row (c), they jointly explain 59% of the variation and three factors maintain their statistical significance while the price-rent ratio loses its marginal predictive power and turns out to be statistically insignificant. Although explaining 59% of the variation in house price growth rates indicates an economically large degree of predictability for future real house prices, adding the price-rent ratio into three-factor model increases R^2 slightly from 57% to 59%, both providing the same degree of predictability. Hence, we can state that the information contained in factors is more than captured by the price-rent ratio. Therefore, we can omit the price-rent ratio from multivariate regressions including three common factors.

Secondly, I consider the second panel of Table 15, which presents the results of predictive regressions for Mexico. As a benchmark, the price-rent ratio on its own is a statistically significant variable explaining 21% of the total variation in real house price growth rate. When we compare the regression results in row (a) and row (b), we can easily conclude that a fourth factor (F10) neither substantially improves the R^2 nor turns out to be statistically significant. Thus, both statistical significance and Bayesian Information Criteria (BIC) suggest that it is sufficient to include only first three factors in predictive regression. Also, adding the price-rent ratio into the four-factor model does not dramatically change the level of predictability of the model while the price-rent ratio maintains its statistical significance as the benchmark.

Next, for South Africa, the benchmark model of price-rent ratio regressing on real house price growth rates reports 27% of R^2 while creating a statistically significant coefficient for the only independent variable, the price-rent ratio. The three-factor model in row (b) presents 41% of R^2 with statistically significant common factors. When we add the fourth factor (F12) into the regression although R^2 improves slightly to 43%, F12 stands as a statistically insignificant variable, which is in line with the results of Bayesian Information Criteria (BIC). Hence, we omit the fourth factor from my regression. Moreover, by comparing row (d) with row (e), we can expect the embedded information in price-rent ratio to be reflected in financial and real variables of South Africa. Hence, we would suggest using three-factor regression in row (b) as the predictive model. Lastly, as a benchmark model of the price-rent ratio for Turkey, row (b) reports statistically significant price-rent ratio explaining 26% of the variation in real house price growth. The three-factor model explains the 61% of the total variation and all the factors are statistically significant. Adding the price-rent ratio into three-factor model provides the same degree of predictability while the price-rent ratio loses its marginal predictive power and becomes a statistically insignificant variable. Hence, we would prefer to use three-factor model.

		$F_{1,t}$	$F_{2,t}$	$F_{3,t}$	$F_{4,t}$	$F_{6,t}$	$F_{10,t}$	$F_{12,t}$	pr_t	R^2
Brazil	(a)		**-0.55**		**0.40**	**0.24**				57%
			[-10.00]		[6.60]	[4.90]				
	(b)								**0.30**	24%
									[4.32]	
	(c)		**-0.52**		**0.39**	**0.23**			0.17	59%
			[-9.01]		[6.45]	[4.72]			[-1.63]	
Mexico	(a)	**-0.23**		**0.26**	**0.20**		-0.10			45%
		[-2.91]		[4.79]	[2.46]		[-1.27]			
	(b)	**-0.23**		**0.26**	**0.21**					44%
		[-2.80]		[4.55]	[2.50]					
	(c)								**0.98**	21%
									[4.63]	
	(d)	**-0.23**		**0.21**	**0.19**		-0.09		**0.64**	46%
		[-3.04]		[3.87]	[3.37]		[-1.12]		[3.09]	
South Africa	(a)		**-0.17**		**0.32**		**0.24**	-0.25		43%
			[-2.12]		[4.66]		[2.44]	[-1.59]		
	(b)		**-0.17**		**0.33**		**0.26**			41%
			[-2.17]		[4.72]		[2.56]			
	(c)								**0.46**	27%
									[5.29]	
	(d)		-0.01		**0.29**		0.16	-0.11	**0.29**	41%
			[-0.04]		[3.85]		[1.72]	[-1.45]	[4.11]	
	(e)				**0.28**				**0.35**	38%
					[3.71]				[4.11]	
Turkey	(a)	**-0.60**	**0.38**		**0.41**					61%
		[-6.73]	[5.09]		[3.48]					
	(b)								**0.43**	26%
									[5.01]	
	(c)	**-0.53**	**0.34**		**0.38**				0.15	62%
		[-4.52]	[4.00]		[3.48]				[0.98]	

Table 15: Single Factor Regression Model $y_{t+1} = \beta_0 + \beta_1 F_t + \beta_2 pr_t$: The table reports estimates from ordinary least square (OLS) regressions of real house price growth rate on the lagged variables named in each column. F_t and pr_t are estimated factors for particular country and price-rent ratio respectively. Newey(1980) corrected t-statistics are reported in brackets. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table.

	Price-rent ratio		Joint Model Specification				
	pr_t	R^2	$F_{2,t}$	$F_{4,t}$	$F_{6,t}$	pr_t	R^2
h=1							
OLS estimate	**0.30**	24%	**-0.52**	**0.39**	**0.23**	0.17	59%
	[4.32]		[-9.01	[6.45]	[4.72]	[1.63]	
h=2							
OLS estimate	**0.28**	19%	**-0.31**	**0.26**	**0.11**	**0.29**	56%
	[2.72]		[-4.56]	[-4.27]	[2.29]	[5.30]	
h=4							
OLS estimate	**0.24**	16%	**-0.30**	**0.22**	**0.10**	**0.29**	54%
	[2.98]		[-4.13]	[-3.48]	[1.98]	[4.95]	
h=8							
OLS estimate	0.15	8%	**0.36**	**0.27**	**-0.14**	0.12	52%
	[1.61]		[6.42]	[4.84]	[3.49]	[1.49]	

Table 16: In-sample results for Brazil: This table reports results of predictive regressions for the h-quarter ahead real house price growth rate using the price-rent ratio model and the joint specification. For each regression, the table reports OLS estimates of the slope coefficients, 90% confidence intervals for the estimates, and the R^2 statistic. Bold font indicates statistical significance.

	Price-rent ratio		Joint Model Specification				
	pr_t	R^2	$F_{1,t}$	$F_{3,t}$	$F_{4,t}$	pr_t	R^2
h=1							
OLS estimate	**0.98**	20%	**-0.22**	**0.19**	**0.18**	**0.73**	44%
	[4.63]		[-2.66]	[3.13]	[2.97]	[3.24]	
h=2							
OLS estimate	**0.74**	12%	**-0.29**	**0.13**	**0.17**	**0.71**	45%
	[1.99]		[-3.10]	[1.98]	[2.02]	[3.31]	
h=4							
OLS estimate	-0.48	8%	**-0.19**	**0.09**	**0.14**	-0.43	41%
	[-1.71]		[-2.02]	[2.19]	[2.28]	[-1.51]	
h=8							
OLS estimate	-0.01	1%	**0.31**	**0.15**	**0.14**	0.15	39%
	[-0.02]		[3.32]	[2.11]	[1.97]	[0.65]	

Table 17: In-sample results for Mexico: This table reports results of predictive regressions for the h-quarter ahead real house price growth rate using the price-rent ratio model and the joint specification. For each regression, the table reports OLS estimates of the slope coefficients, 90% confidence intervals for the estimates, and the R^2 statistic. Bold font indicates statistical significance.

	Price-rent ratio		Joint Model Specification				
	pr_t	R^2	$F_{2,t}$	$F_{4,t}$	$F_{10,t}$	pr_t	R^2
h=1							
OLS estimate	**0.46**	27%	**-0.28**	**0.25**	**0.21**	**0.31**	47%
	[5.29]		[-3.77]	[2.54]	[2.32]	[3.52]	
h=2							
OLS estimate	0.26	8%	**-0.53**	**0.38**	**0.20**	0.15	45%
	[1.68]		[-2.07]	[4.81]	[2.03]	[1.66]	
h=4							
OLS estimate	-0.23	4%	**-0.27**	**0.42**	0.16	-0.06	38%
	[-1.60]		[-2.05]	[5.30]	[1.67]	[-0.69]	
h=8							
OLS estimate	-0.07	1%	**-0.26**	**0.32**	**0.24**	-0.05	40%
	[0.76]		[-2.20]	[3.44]	[2.38]	[-0.66]	

Table 18: In-sample results for S.Africa: This table reports results of predictive regressions for the h-quarter ahead real house price growth rate using the price-rent ratio model and the joint specification. For each regression, the table reports OLS estimates of the slope coefficients, 90% confidence intervals for the estimates, and the R^2 statistic. Bold font indicates statistical significance.

	Price-rent ratio		Joint Model Specification				
	pr_t	R^2	$F_{1,t}$	$F_{2,t}$	$F_{4,t}$	pr_t	R^2
h=1							
OLS estimate	**0.45**	26%	**-0.53**	**0.34**	**0.38**	0.15	55%
	[3.86]		[-4.52]	[4.00]	[3.48]	[0.98]	
h=2							
OLS estimate	**0.34**	15%	**-0.35**	**0.20**	**0.52**	**0.22**	58%
	[2.69]		[-2.41]	[1.96]	[3.99]	[2.05]	
h=4							
OLS estimate	**0.31**	19%	**0.27**	**0.32**	**0.50**	**0.28**	51%
	[3.93]		[-2.05]	[3.20]	[3.84]	[3.02]	
h=8							
OLS estimate	0.22	5%	**-0.66**	**0.25**	**0.36**	0.20	49%
	[1.74]		[-3.46]	[2.27]	[2.53]	[1.86]	

Table 19: In-sample results for Turkey: This table reports results of predictive regressions for the h-quarter ahead real house price growth rate using the price-rent ratio model and the joint specification. For each regression, the table reports OLS estimates of the slope coefficients, 90% confidence intervals for the estimates, and the R^2 statistic. Bold font indicates statistical significance.

2.6.3.1 *Interpretation of Factors*

One of the main outcomes of Table 15 is that the statistical significance test results and Bayesian Information Criteria (BIC) suggest the same level of three factors to be included in our predictive regressions and forecast specifications. Moreover, even though the price-rent ratio is a statistically significant variable in explaining variation in real house prices growths on its own, adding it into the three-factor model does not improve the predictability substantially. The estimated three-factor models have statistically and economically significant predictive power beyond that contained in price-rent ratio. Next, I relate the factor interpretations to the sign of the slope coefficients of three-factor model predictive regressions. One of the most impressive results of this research is the mutuality of top three factors predicting the real house price fluctuations in four emerging countries; Brazil, Mexico, South Africa, and Turkey. The first-factor component, which is mutual in all four emerging countries and represented by $F_{BR,2}$ for Brazil, $F_{MX,1}$ for Mexico, $F_{SA,2}$ for South Africa, and $F_{TR,1}$ for Turkey, is related to financial variables. Financial variables factor has negative coefficients in all of the regressions for each country. The negative sign of the coefficient suggests that expected house price growth rates move countercyclical since lower financing rates lead to an increase in house demand due to cheaper mortgage loans, which in result boosts the house prices. This finding is in line with [87] who document that large capital flows to emerging countries led to sharp declines in bond yields which fuelled a house price boom emerging countries.

The second-factor component is **monetary and credit quantity aggregates**, which is again all observed in house markets of Brazil, Mexico, South Africa, and Turkey. Monetary and credit quantity aggregates factor has a positive slope coefficient in all of the regressions for each country. The positive sign of the coefficient suggests that expected house price growth rates move procyclical. Since the housing market is both consumption and investment good, it is reasonable to expect house prices to rise as the demand for housing expands due to an increase in money supply. This finding show why regulators and central bankers of emerging market countries more actively monitor monetary and credit quantity aggregates such as loan growth and outstanding credit and use macro-prudential tools (see [88]) to prevent potential financial instability coming from booming mortgage markets fuelled house prices when necessary. In the following section, I discuss in detail how policymakers in emerging market countries react to rising threat to financial stability caused by rising house prices beyond fundamental reality. The last factor component for each of the countries is **real economic activity**, which has a positive slope coefficient in all of the regressions for each country. The positive sign of the coefficient suggests that predicted house price growth rates move procyclical. When there is an economic boom, the affordability of households improves; hence we can expect to see a positive impact on future house price growth rates, and vice versa. [69] also show that employment and income variables have a positive relation with future housing returns. The differential behavior of house prices across countries that I focus on here constitutes a tangible manifestation of those real economic differences. This divergence has significant implications for portfolio managers. The non-systematic, diversifiable, risk of a portfolio can be reduced with a global portfolio strategy by focusing on country-specific factors including employment, income and price levels.

The price-rent ratio is known as the most widely used housing market predictor. In Table 16 to 19, I compare the forecasting power of joint model of lagged price-rent-ratio and the lagged factors with that of the price-rent ratio for Brazil, Mexico, South Africa, and Turkey, respectively.Table 16 to 19 show that the price-rent ratio on its own has less predictive power than joint factor models at all forecasting horizons with an R^2 level less than 30%. By including the three factors, the predictive power as measured by the R^2 increases substantially across all forecasting horizons. The three factors are generally statistically significant across horizons, while the price-rent ratios significance level in joint model depends on its significance level in the stand-alone model. In other terms, if the price-rent ratio is insignificant on its own, we can expect it to remain insignificant in joint regression with factors. On the other hand, it might turn out to be insignificant in joint regression even though it is significant stand-alone. Hence, my results suggest that it is insufficient and misleading to base house price forecasts on a single predictor of price-rent ratio only. Despite its popularity, the price-rent ratio is not a comprehensive variable that represents all relevant information. In the in-sample predictions, I estimate time-t factors using a full data sample, instead of using data only up to time-t. In this way, I obtain efficient estimates of latent factors that more accurately represent the covariance structure of the large panel of variables, since I do not discard the information included in the full sample. Further, the main aim of in-sample analysis is not a real-time prediction but rather an accurate estimation of the predictive relation between real house price movements and common factors. Finally, I also address the important issue of real-time prediction through an out-of-sample predictive exercise, where I re-estimate the factors recursively each period.

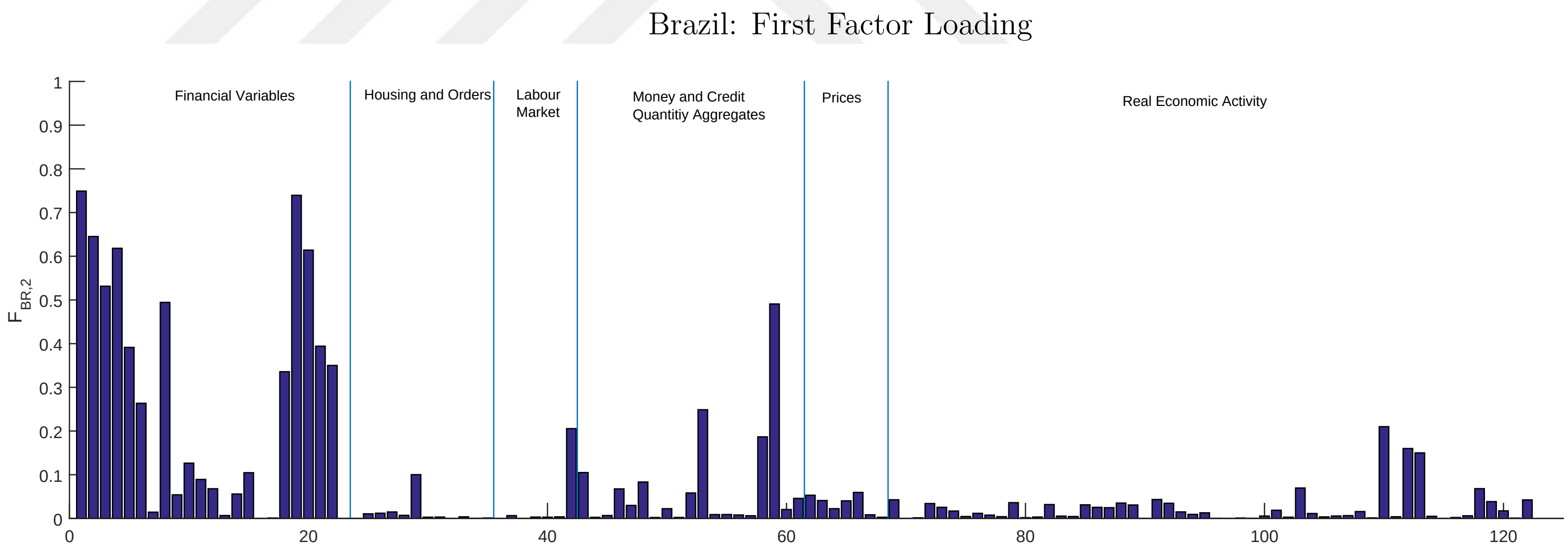


Figure 19: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{2},\mathbf{BR}}$ for Brazil. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

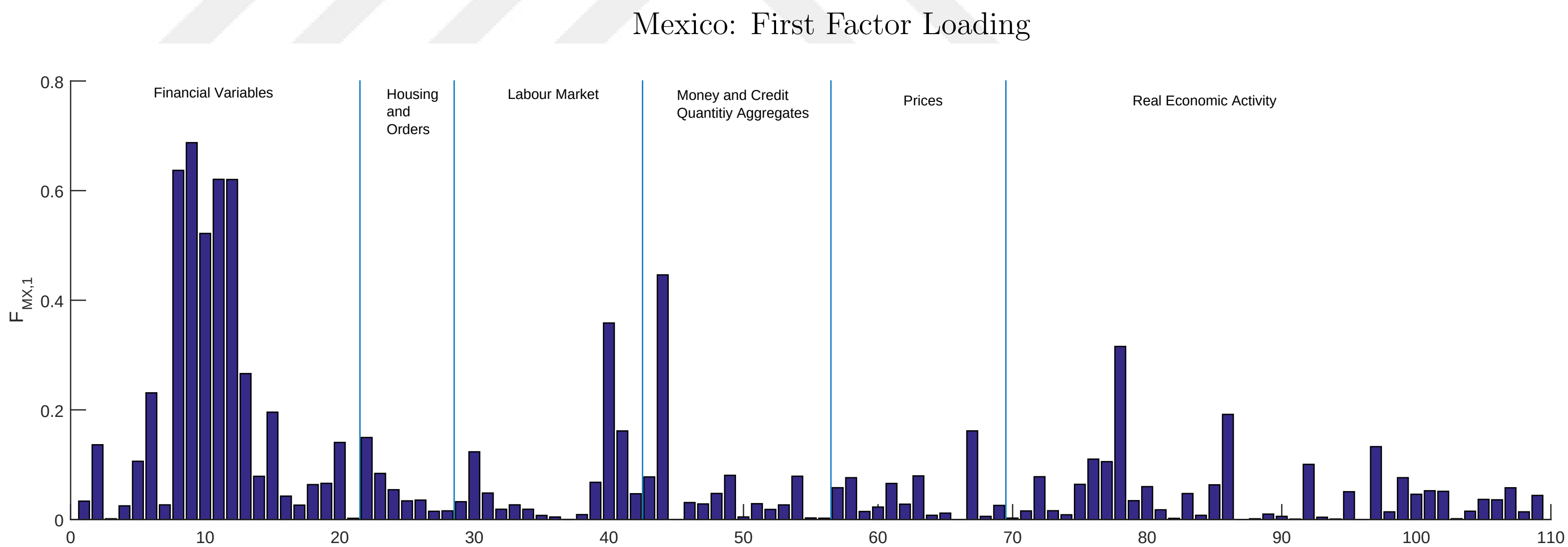


Figure 20: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{1},\mathbf{MX}}$ for Mexico. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

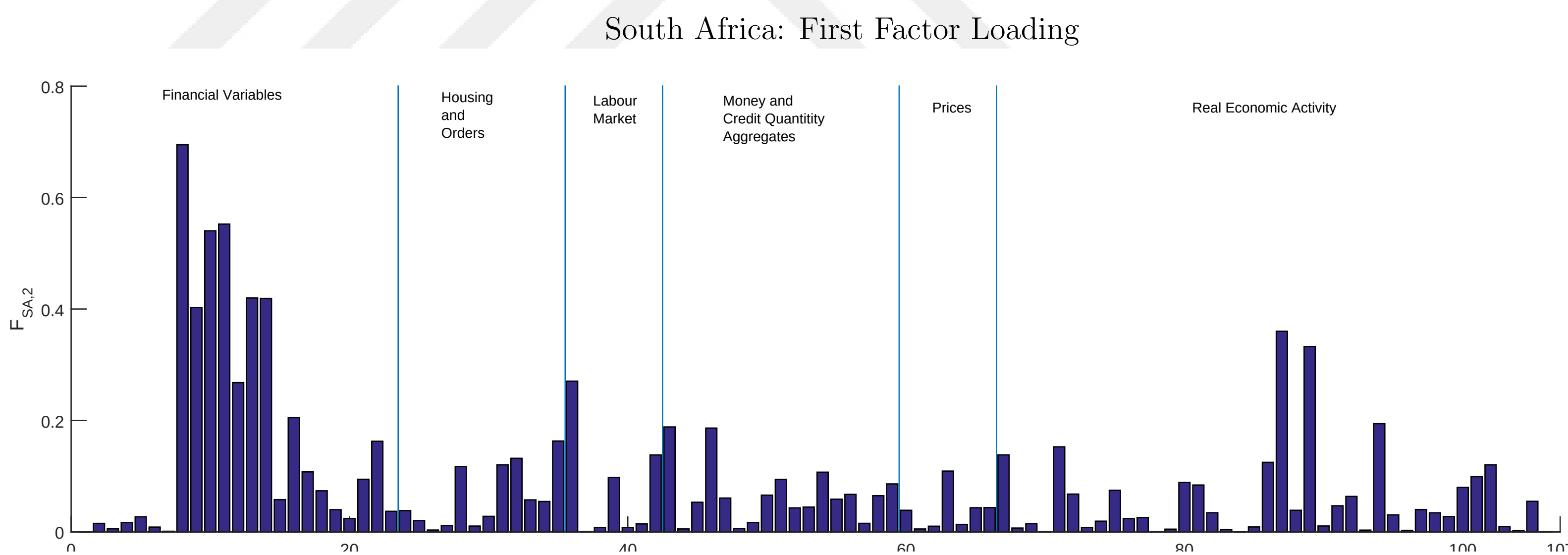


Figure 21: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{2},\mathbf{SA}}$ for South Africa. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

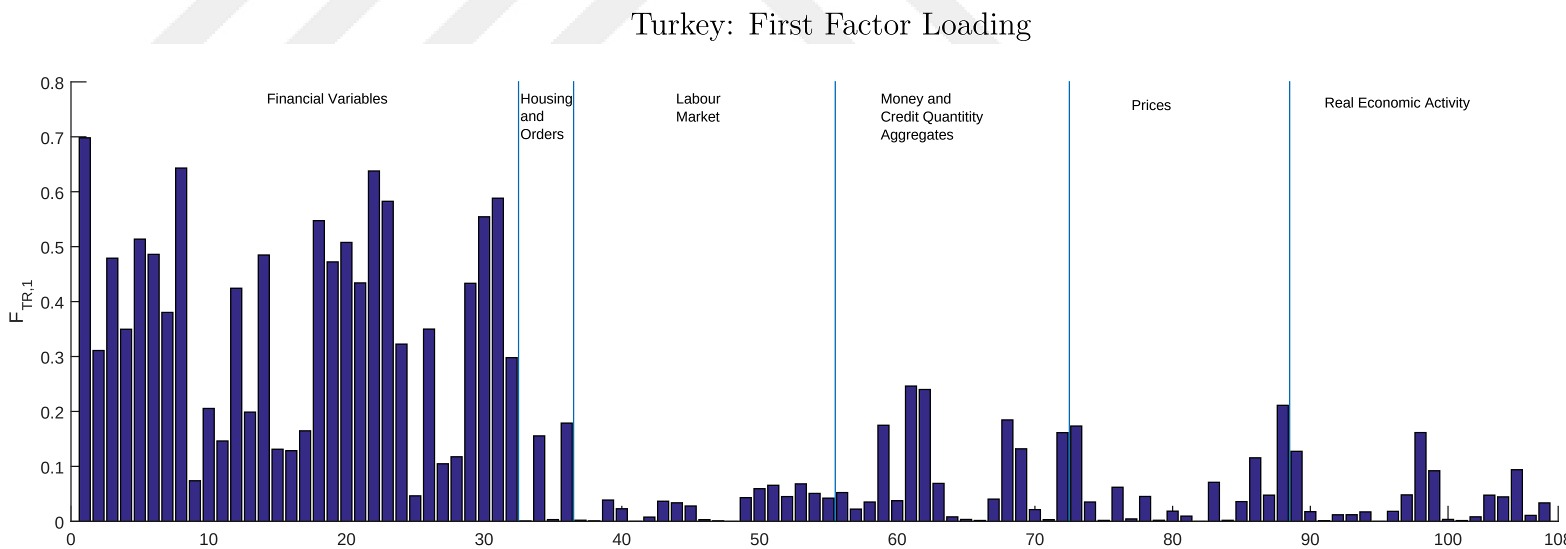


Figure 22: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto first factor loading, $F_{\mathbf{1},\mathbf{TR}}$ for Turkey. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

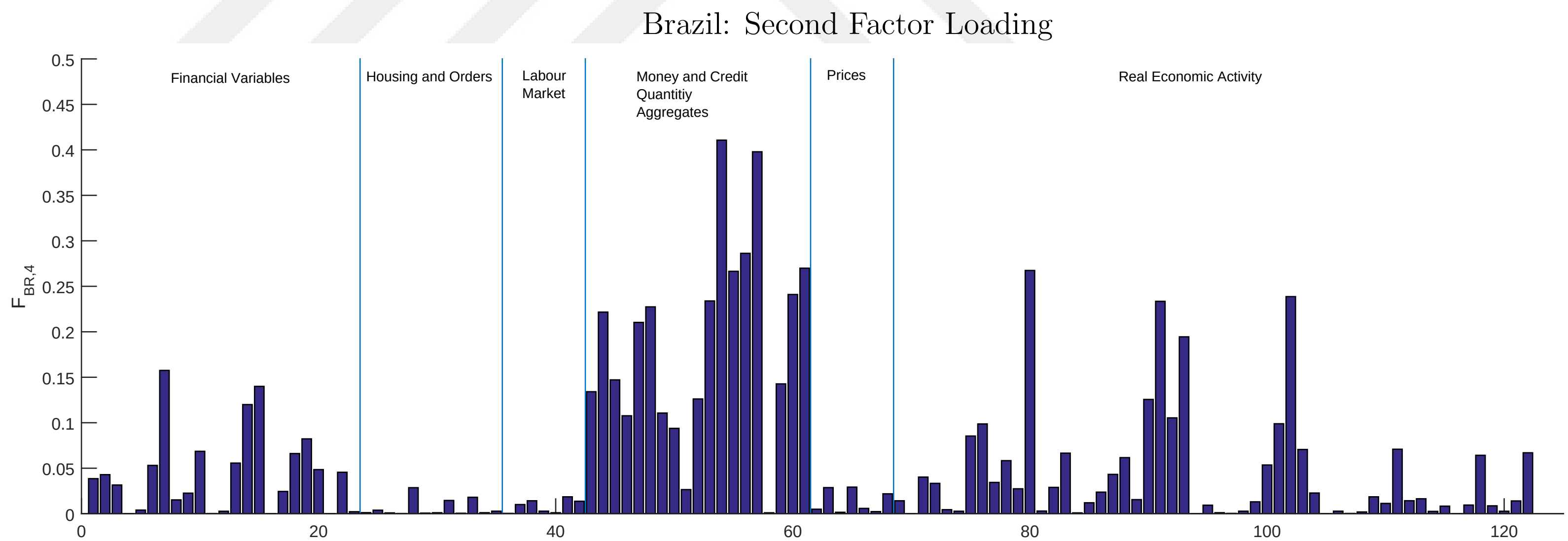


Figure 23: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{4,BR}}$ for Brazil. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

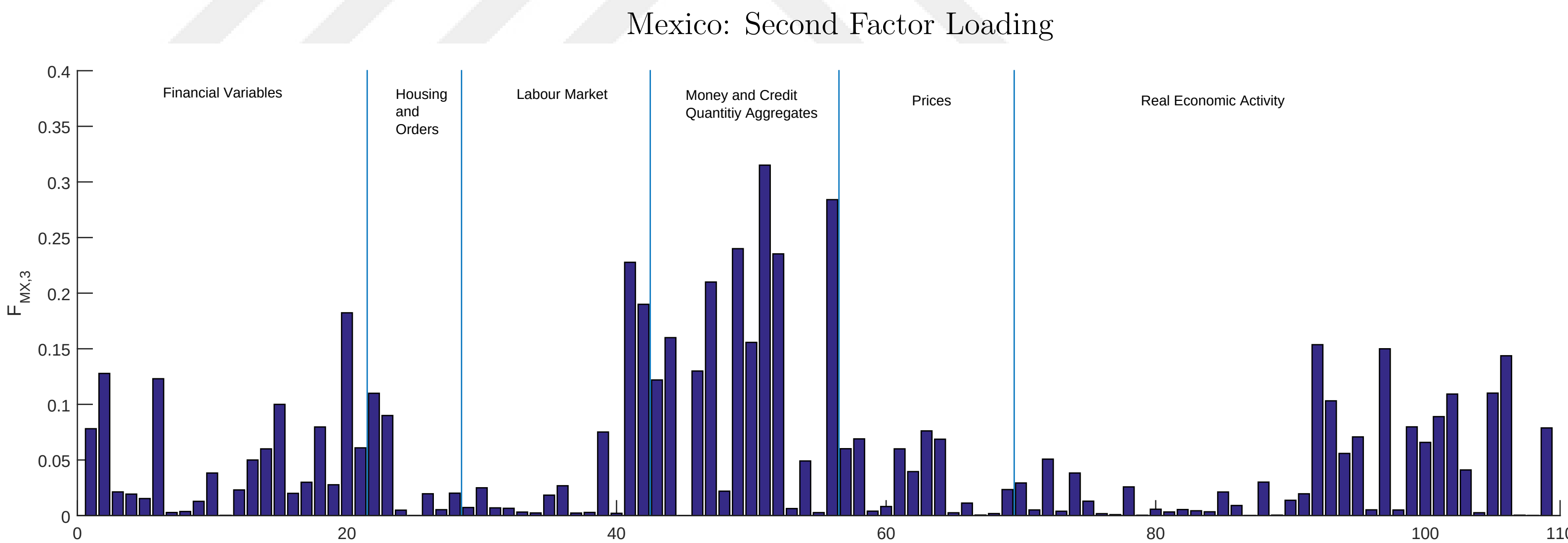


Figure 24: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{3,MX}}$ for Mexico. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

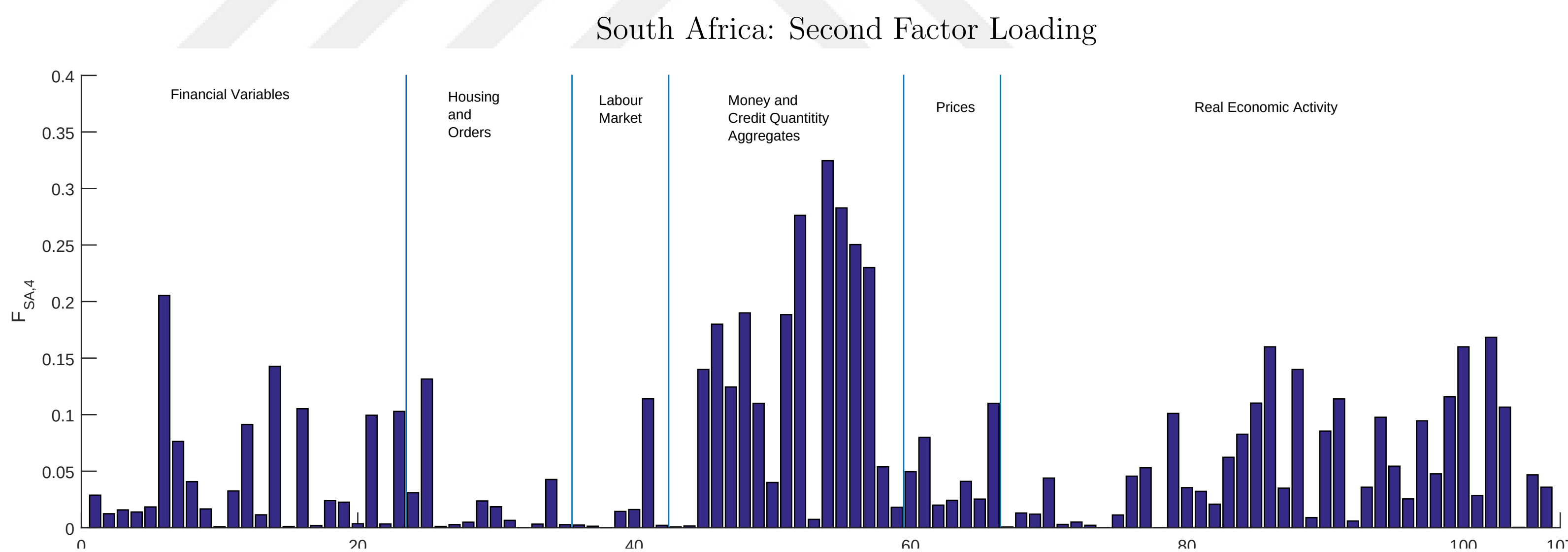


Figure 25: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{4,SA}}$ for South Africa. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

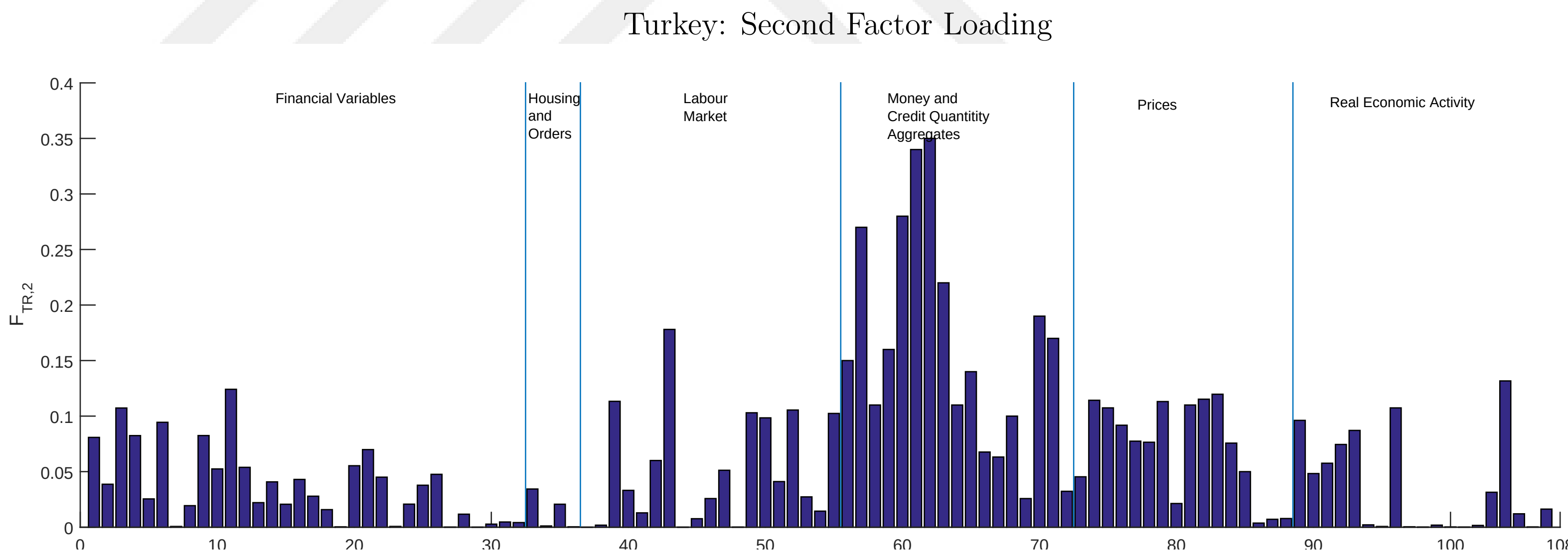


Figure 26: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto second factor loading, $F_{\mathbf{2},\mathbf{TR}}$ for Turkey. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

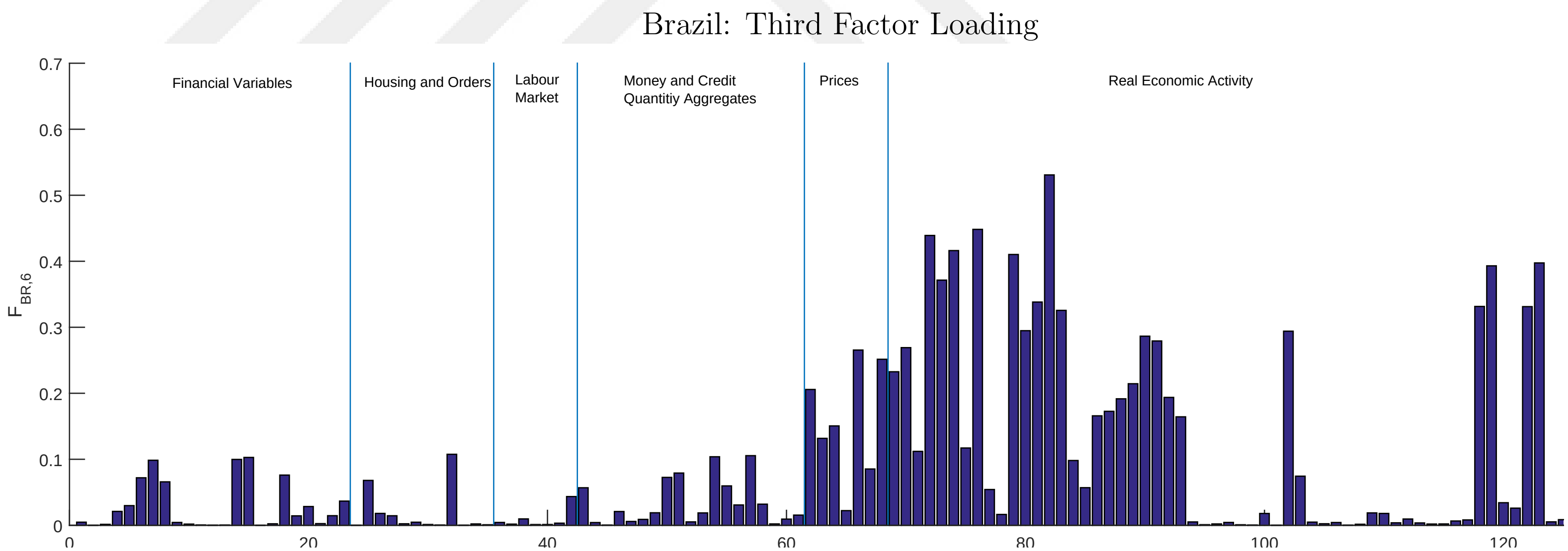


Figure 27: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto third factor loading, $F_{\mathbf{6},\mathbf{BR}}$ for Brazil. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

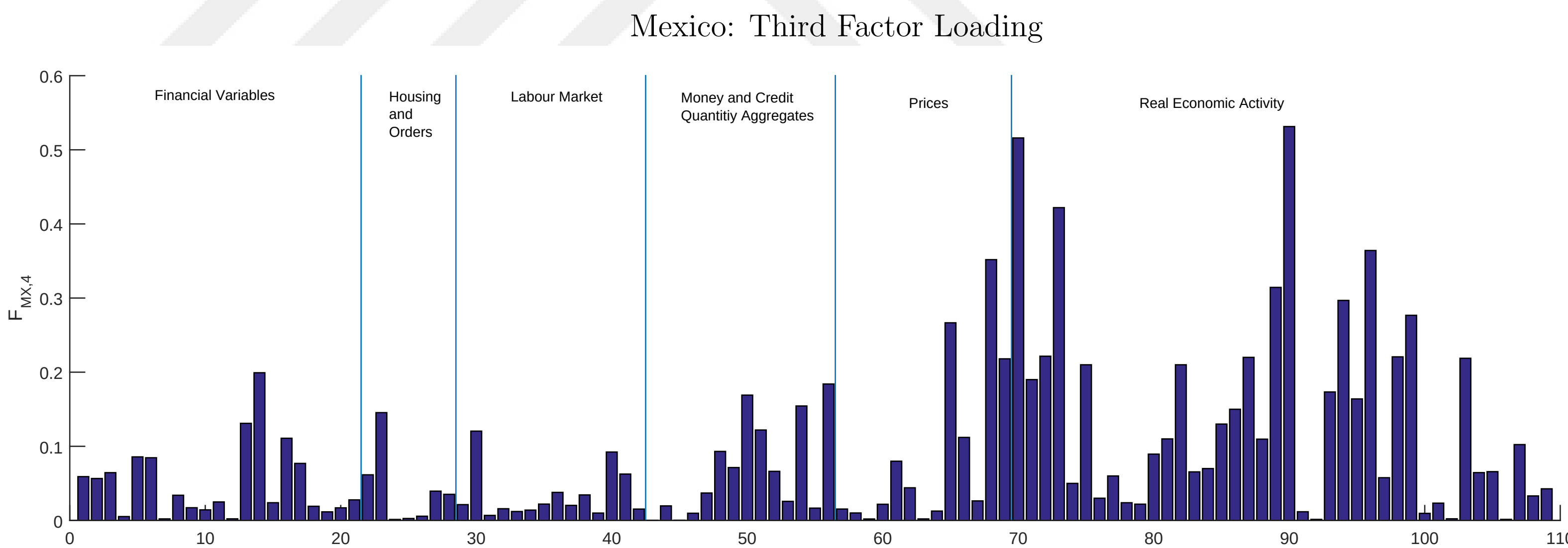


Figure 28: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto third factor loading, $F_{\mathbf{4,MX}}$ for Mexico. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

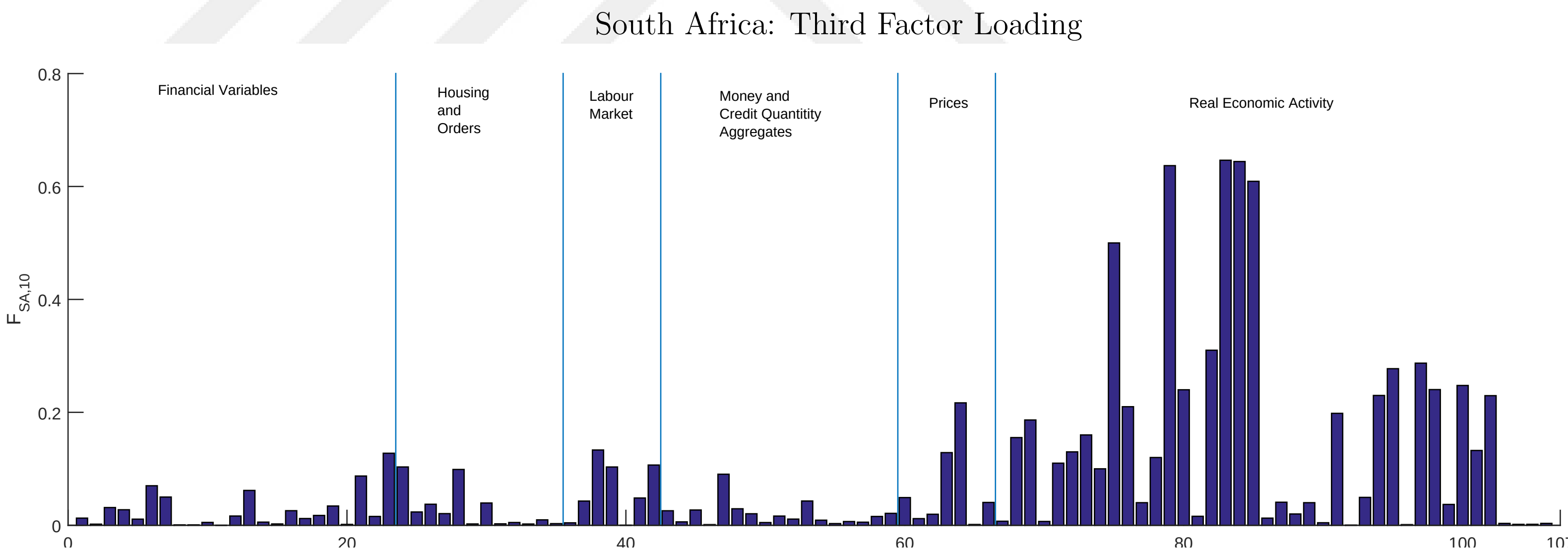


Figure 29: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto third factor loading, $F_{\mathbf{10,SA}}$ for South Africa. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

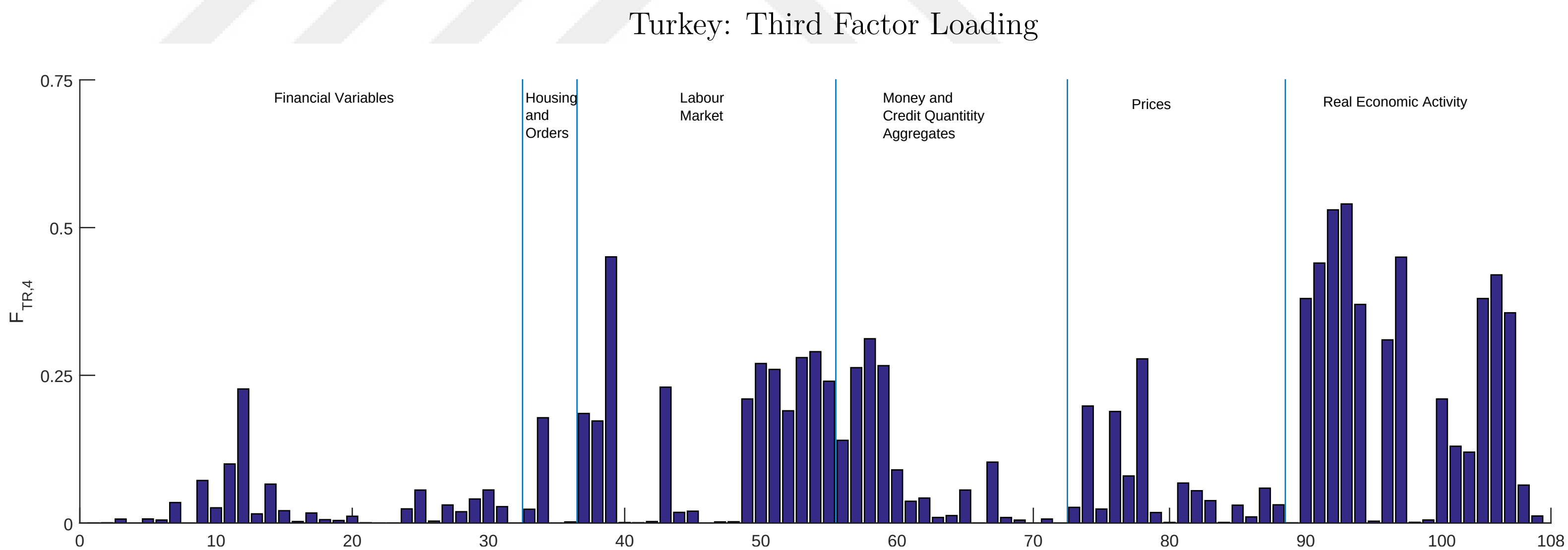


Figure 30: The chart displays the R^2 statistics as bar charts by regressing the series number given on the x-axis onto third factor loading, $F_{\mathbf{4,TR}}$ for Turkey. The time series are categorized into six subgroups: Real Economic Activity, Labour Market, Housing and Orders, Prices, Money and Credit Quantity Aggregates, Financial Variables. The subgroups are separated by lines.

2.6.4 Out-of-sample

I now turn to the question of whether common factors have predictive ability for real house price growth in an out-of-sample exercise. The main point in an out-of-sample prediction exercise is that the forecast for the time t + 1 can be made using the data available only up to time t. In the first step, I use a large panel of macroeconomic and financial data and estimate the common factors up to time t.

In Table 20, I use the Mean Squared Forecast Error (MSFE) to measure the out-of-sample performance of the models. The out-of-sample forecasting exercise uses an initial estimation period of 22 quarters from 2007:Q2 to 2012:Q4, and the out-of-sample period thus runs from 2012:Q4 to 2015:Q2.

I test whether the three-factor model produces statistically significant reductions in the MSFE relative to the benchmark models. Table 20 reports the MSFE-ratio between the three-factor model and each of the three benchmark models. The table shows that across all forecasting horizons, h = 1; 2; 4; 8, the three-factor model yields lower MSFE values than the historical mean and autoregressive benchmarks for Brazil and Turkey. The three-factor model substantially outperforms the price-rent ratio model for Brazil and Turkey.

The strong predictive power of my three-factor model suggests that it is not only insufficient but also misleading to form house price forecasts based on a limited set of variables. I find that the price-rent ratio, one of the most widely used house price indicators, performs worse than the factor model in both In-sample and out-of-sample forecasting models. Besides, model strongly beats the historical mean, but also performs remarkably well compared to both an autoregressive benchmark with a rich lag structure as well as to computationally intensive factor forecast combination models.

	Model	Benchmark	horizon h=1	h=2	h=4	h=8
Brazil	Three Factor Model	Price-Rent Ratio	0.325	0.19	0.311	0.785
	Three Factor Model	Mean	0.345	0.177	0.112	0.290
	Three Factor Model	AR1	0.952	0.208	0.269	0.689
Mexico	Three Factor Model	Price-Rent Ratio	1.037	0.924	1.048	0.917
	Three Factor Model	Mean	1.065	0.998	0.855	0.927
	Three Factor Model	AR1	1.046	0.893	0.887	0.918
South Africa	Three Factor Model	Price-Rent Ratio	0.805	0.980	1.097	0.868
	Three Factor Model	Mean	0.763	0.953	1.011	0.842
	Three Factor Model	AR1	0.563	0.980	1.051	0.835
Turkey	Three Factor Model	Price-Rent Ratio	0.809	0.826	0.868	0.910
	Three Factor Model	Mean	0.658	0.512	0.550	0.833
	Three Factor Model	AR1	0.777	0.629	0.597	0.951

Table 20: Out-of-sample results: This table reports the ratio of the mean squared forecast error (MSFE) between the three-factor model and various benchmarks. t-statistics are given in parenthesis below the MSFE ratios. Following [77], I report the Diebold Mariano t-statistic when comparing the three-factor model with the price-rent ratio model and when comparing the three-factor model with the historical mean and autoregressive models, I report the Clark-West t-statistic. The null hypothesis of equal MSFE is rejected when the t-statistic is greater than 1.28 (one sided test using 10% significance level). Bold font indicates statistical significance. The out-of-sample window is 2012:Q4 to 2015:Q2.

2.6.5 Extracted Factors and Macro-prudential Housing Policies

In the aftermath of global financial crisis, discussions around global housing markets have turned their focus from shrinking prices in developed market especially in the United States and Western European countries to increasing worries about the potential for bubbles emerging in some developed and emerging market countries

Existing literature has already investigated the likelihood for countries with substantial house price increases to experience a significant subsequent downturn which could potentially harm overall macroeconomic conditions. ([66], [65] and [67]) Over the past years, many advanced and developing countries including the ones which I investigate in this study have seen substantial increases in house prices. (see: Figure 31) A more surprising outcome from this figure is that the countries with largest real house prices, including Sweden, UK, Israel and Turkey, did not experience significant downturns during the global financial crisis. This replaced the concerns around housing market bust with concerns about housing bubbles and its potential negative consequences on financial stability. To address these systemic concerns, central banks and other regulators around the globe have taken macroprudential measures to curb expanding mortgage credit and rising house prices. The term macroprudential comes from combining two words: Macro to emphasize that the policy should be implemented by either a governmental or a regulatory body and Prudential, which means that the policy action is aimed at preventing a potential damage on financial stability or the broader economy in a countercyclical manner. Acting countercyclical is critical as cyclical policy responses may not address the underlying problem sufficiently.

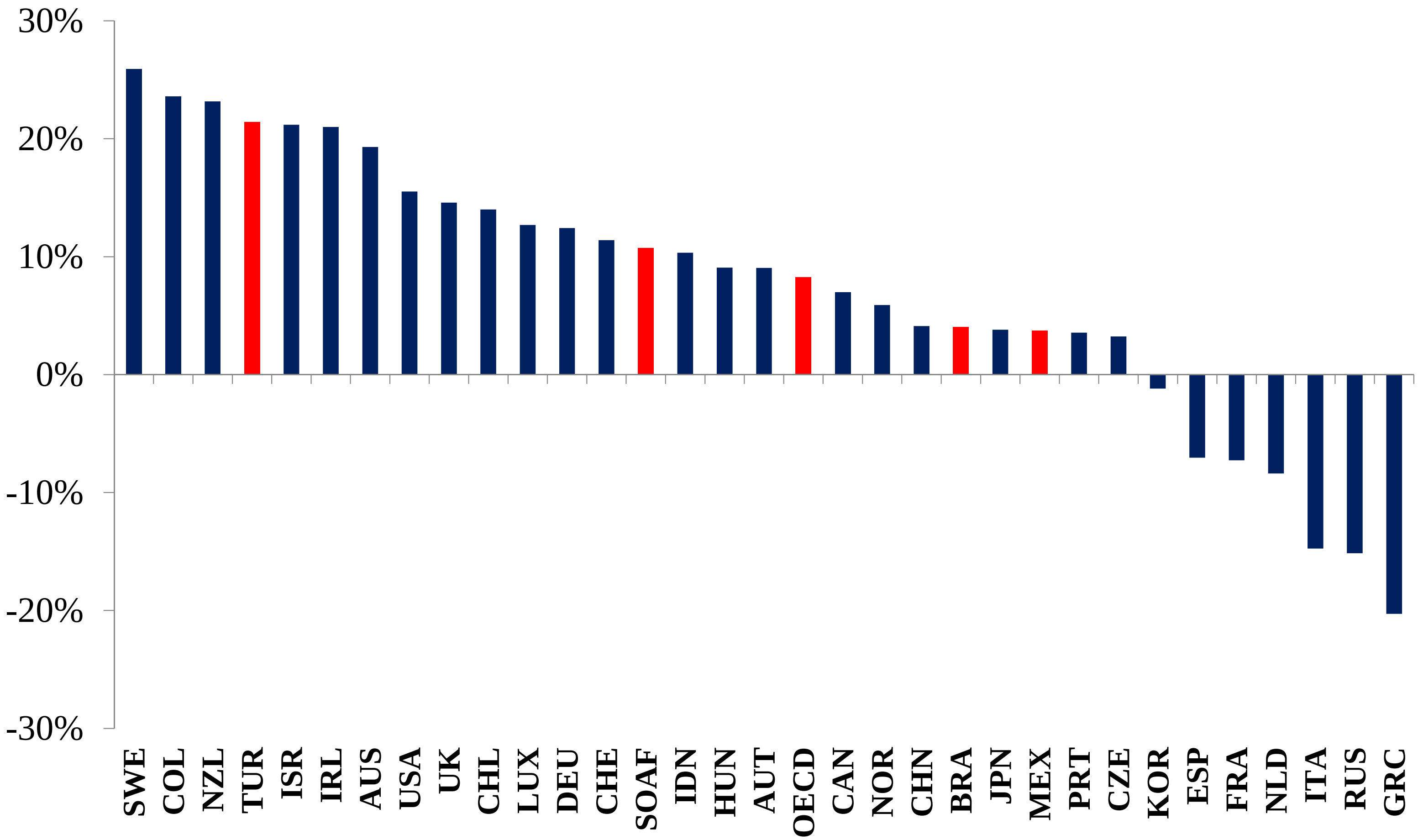


Figure 31: Cumulative Real House Price Changes in OECD Countries (2012-2015)

In a study to measure effectiveness of macroprudential to prevent housing price bubbles from the making, [88] documents that financial authorities in many advanced and emerging market economies have actively used macroprudential policy tools as a response to rising house prices. Most of these policy instruments aimed at curbing cyclical effects of easy monetary policies implemented by global central banks which aimed at reviving economic activity after the global financial crisis. This proactive approach shows why macroprudential measures which target the housing sector are usually implemented in tandem with central bank policy activity such as changes in lending rates and reserve requirements and capital flow management measures. Using a large panel dataset based on Database for Policy Actions on Housing Markets report published by BIS and IMF surveys, [89] finds that three types of macroprudential tools have the most significant impact on housing credit growth. These tools include debt-service-to-income (DTI) ratio, the maximum loan-to-value (LTV) ratio, and housing-related taxes or capital requirement.

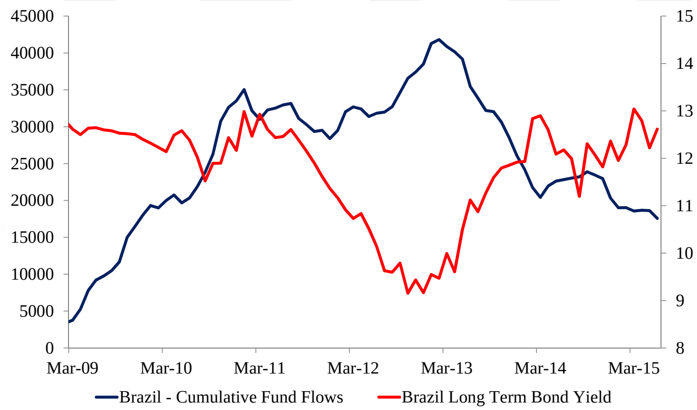
In empirical results sections, I showed that Financial variables factor has negative coefficients in all of the regressions for each country which means that lower interest rates lead to an increase in house demand due to cheaper mortgage loans, which in result boosts the house prices. Central Banks including FED, ECB and BOJ implemented ultra-loose monetary policies which included the purchase of long-dated government debt, investment grade corporate bonds and agency bonds, which resulted in significant drop in nominal interest rates on long-term debt instruments globally. ([90]) Yields on emerging market debt also fell on the back of this sustained monetary stimulus, and this was also reflected in local banks borrowing costs which in turn reduced the mortgage lending rates.

Figure 32 shows the impact of cumulative fund flows to emerging markets and their impact on long-term bond yields. All countries in my analysis attracted large inflows after the global financial crisis as central banks stepped in with aggressive monetary tools. Long-term bond yields declined substantially in all countries until FED tapering announcement in May 2013. Moreover, I also empirically showed that monetary and credit quantity aggregates factor has a positive sign of the coefficient which suggests that expected house price growth rates move procyclical, and house prices tend to rise in tandem with an increase in money supply. Figure 33 shows the impact of falling interest rates on outstanding mortgage loans.

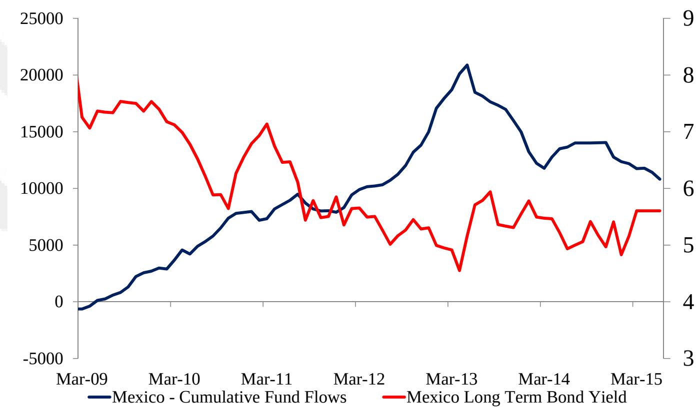
In all countries but S.Africa, real outstanding mortgage loan stock grew enormously driven by fund flows and drop in yields. Due to a very low base of mortgage indebtedness, the fastest real loan growth was observed in Brazil and Turkey (90% and 120%) where the ratio of outstanding mortgage debt to GDP was only 2.2% and 3.6% respectively in 2008. This compares to 9.4% mortgage debt to GDP in Mexico where real loan growth was about 40% during the same period. In South Africa, mortgage debt to GDP was 42.3% in 2008 and mortgage loan volume contracted in real terms as households preferred to reduce their existing mortgage debt rather than investing in new residential properties. According to BIS statistics for 2015, even after sharp increases in mortgage stock, current mortgage debt to GDP in Brazil and Turkey (3.6% and 6.8%) is still well below advanced economies average (61.4%) and emerging markets average (10.6%).

However, due to the bad experiences in US Housing market crisis in 2008 and European Debt Crisis in 2011, regulators and central bankers of emerging market countries closely monitor monetary and credit quantity aggregates such as loan growth and outstanding credit and react more proactively to use macro-prudential tools. When we look at the pre-taper period (March 2009-May 2013) where the credit expansion was fastest, we see that all countries in my analysis but South Africa enacted countercyclical policies to control credit growth fuelled by portfolio flows and to prevent an economy from overheating. Among these countries, Turkey and Mexico specifically targeted housing sector, and Turkish Banking Regulator introduced 75% LTV ratio limit for housing loans and Mexican authorities tightened lending standards for new mortgage lending.

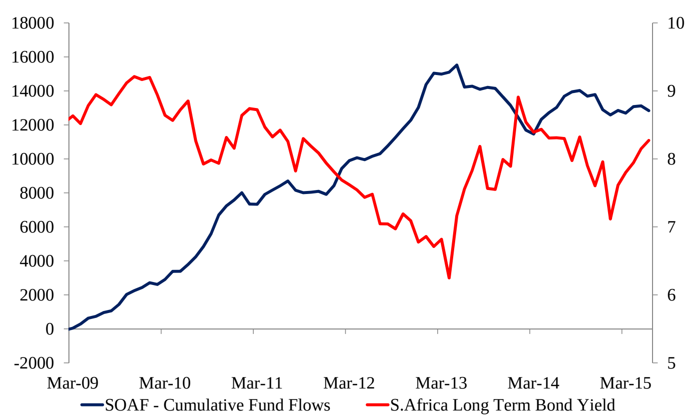
Identifying the cause and effect relationship between macro-prudential measures and housing prices in emerging markets is a difficult task as house prices continued to rise in many areas in the world despite tightening in global financial conditions and increasing the cost of financing in emerging market economies. The real price increase in housing can go beyond fundamental reality if the underlying driver is related to expectations on future price increases or safe haven demand.



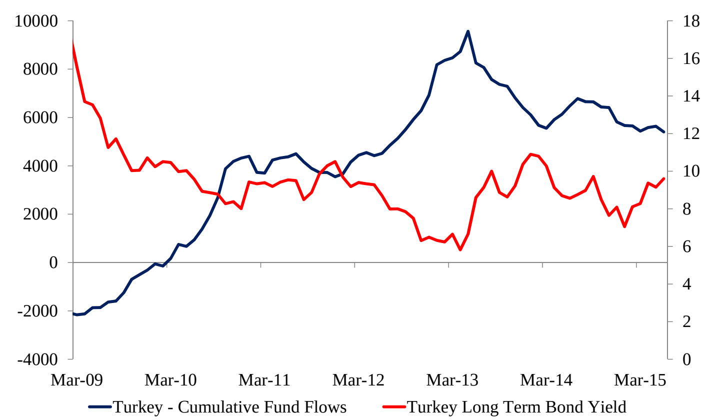
(a) Brazil



(b) Mexico

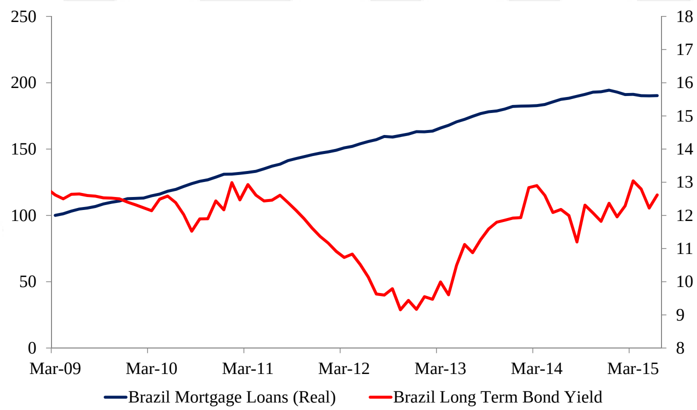


(c) South Africa

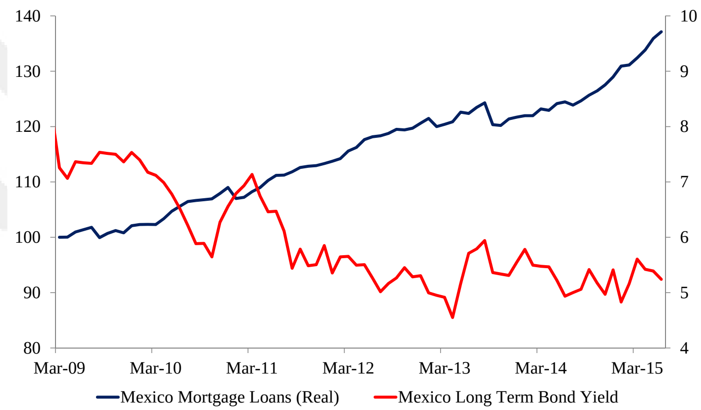


(d) Turkey

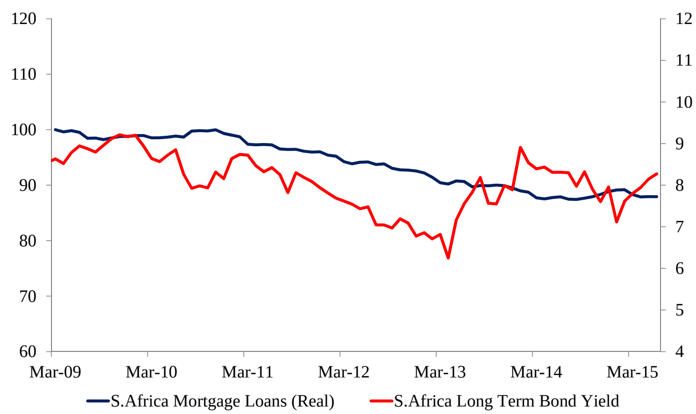
Figure 32: Cumulative Fund Flows and Long Term Yields. Cumulative fund flows are shown in million $ terms. EPFR data is used to measure cumulative fund flows to emerging economies starting from June 2007. For long term bond yields, 5 to 10 Year maturities are used.



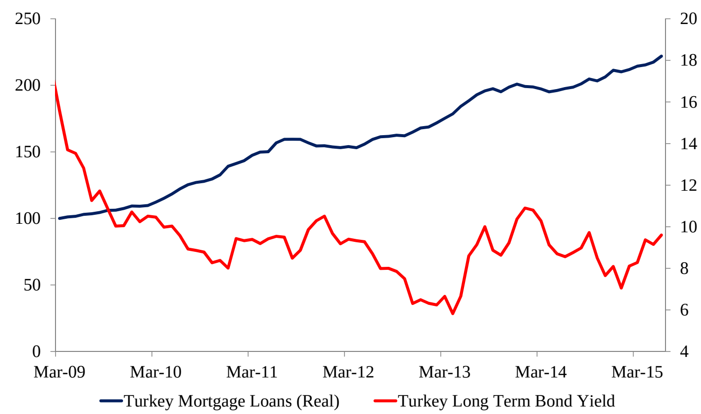
(a) Brazil



(b) Mexico



(c) South Africa



(d) Turkey

Figure 33: Long Term Yields and Mortgage Loan Stock. Graphs show deflated outstanding mortgage loans for each country. I normalize outstanding loans at 100 starting at 31/03/2009. I use national data sources for outstanding mortgage loans and to deflate data I use CPI indices for each country as provided by Bloomberg. For Long term bond yields, 5 to 10 Year maturities are used.

2.7 Conclusion

In this chapter, I contribute to the existing empirical literature on housing prices by showing that macroeconomic fundamentals have significantly important predictive power for housing markets of four major emerging market countries: Brazil, Mexico, South Africa, and Turkey, over the period of 2007 Q2 to 2015 Q4. I use a dynamic factor model based on a set of common factors that are extracted from a large panel of macroeconomic data. I exploit information from large economic and financial time series to assess: (i) the degree to which a small number of statistical factors, regardless of their nature, can be used to understand a broad set of economic indicators; (ii) the degree to which the estimated factors, identified statistically, relate back to the set of macroeconomic variables; and (iii) the degree to which the estimated factors can predict real house price growth rates.

I emphasize two aspects of my findings. First of all, I find that a small set of three key factors could explain a significant portion of the variation in real house price growth rates of sample emerging countries. The R^2s for the regressions are fairly high, indicating that the estimated factors capture much of the variation in real house price fluctuations. Secondly, by relating the estimated factors back to macroeconomic indicators, I find that, in my sample of emerging markets, the first factor reflects the financial variables whereas the second and third factors reflect monetary and credit aggregates and real economic activity, respectively. I contribute to the literature by showing the mutuality of top three factors predicting the real house price fluctuations. Furthermore, I provide the evidence that the predictive power of estimated factors is not just statistically significant, but also economically important. My results also highlight that it is important to include data rich macroeconomic factors when forecasting house prices. The strong degree of predictability that I document using my panel approach suggests that it is insufficient and misleading to form house price forecasts based on a limited set of economic time series. As an illustration of this point, I find that the predictive power of the three-factor model performs better than the price-rent ratio, which is one of the most widely used house price indicators. My research could be furthered with a deeper analysis of housing price synchronization in emerging countries. Although I provide evidence about the sources of common movements in house prices, I lack in articulating why house prices have become more synchronized over time. A natural next topic to explore could be a deeper analysis of differential effects of shocks and structural features of countries, including their linkages through the banking system, on the temporal changes in the degree of synchronization of house prices.

CHAPTER III

FORECASTING TURKISH REAL GDP USING TARGETED PREDICTORS

3.1 Introduction

It is widely known that the information about the current state of economic activity and the forecasting of its short-term prospects are of fundamental importance for policy makers in governments, central banks, and financial markets, as well as the public. Many key statistics are released at low frequency and with long publication delays. Gross domestic product (GDP) figures, a key statistic describing the overall economic performance, are subject to substantial publication lags. Unlike price and financial variable time series which are collected at a higher frequency and published in a more timely manner, an initial estimate of quarterly real GDP is released about two months after the end of each quarter in Turkey. Due to the lack of timely information, policy institutions such as central banks and ministries, are always forced to conduct their policies without knowing current economic performance. In particular, these time lags undermine the ability to make appropriate changes in monetary policy. Hence, it is crucial to obtain an early estimate of current quarterly GDP which is called as "Nowcasting." [1]

In the context of growing data availability, economists and forecasters have access to flow of information from a wide range of macroeconomic data published on a different frequencies, including monthly (e.g. capacity utilization, industrial production, PMI surveys), weekly (e.g. unemployment, money supply), and continuously

[1] Predictions for the current, previous and following quarter are called as nowcast, backcast and forecast respectively (see for example [91].

(e.g. financial variables). To benefit from this tremendous amount of information, factor models have emerged as an exciting alternative for the short term forecasting of quarterly GDP as they efficiently summarize the information contained in large databases. Especially, dynamic factor models have received increasing attention in the literature and been widely used in central banks and research institutions to predict economic variables and conduct monetary policies. (See, [57] [28] for USA; [39] for UK, [92] for the Euro area, [93] for Germany; [94] for Norway; [95] for France; [96] for Italy.) The reason why dynamic factor models have become so popular is that they allow extracting common factors from a large dataset of potential indicators while efficiently handling with the general pattern of missing or mixed frequency data and non-synchronous data releases.

Despite the increasing attention for factor models based on a large dataset, there is no consensus on the selection criteria of variables that would be included in the panel. Also, how do we determine the optimal size of the dataset that we should include in factor analysis? Is there any gain from choosing a limited number of variables? Answers to these questions have particular importance for practitioners and policy makers. Since [83] have proved that factor estimates are consistent for large N and T, the natural choice would be to use as much data as possible for the estimation of factors. The common view among practitioners is that a large number of indicators may improve the forecasts of macroeconomic variables due to containing additional information about the state the of the economy. Also, using numerous variables further reflects a central bankś motivation to take all potentially relevant factor into account ([42]). However, [97] and [98] show that the forecast accuracy of the model does not necessarily improve if the additional series are noisy or unrelated to with the variable to be forecasted. Moreover, [99] indicates that tests and criteria for determining the number of factors are inconsistent when the dataset is not correctly specified. Hence, the forecast accuracy of nowcasting models is highly sensitive to the variable selection procedure.

The current literature suggests different methods to address the sample selection issue. The first stream of sample selection proposes statistical methods. Specifically, [98] suggest eliminating redundant series that are not informative for forecasting the target variable and identifying an efficient set of predictors with LASSO algorithm. The second stream suggests choosing variables with subjective judgment. The size of datasets in the empirical literature varies from 80 to more than 400 variables. Finally, an alternative strategy recommends selecting the variables that are mostly tracked by market analyst. The underlying reason behind this approach is that market participants obsessively monitor all macroeconomic data to get a view on current and future fundamentals of the economy ([100] and [101])

Existing studies investigating the Turkish GDP forecasting are quite limited. [102] implement a small scale factor model to produce Turkish GDP nowcast. They find that using soft indicators such as Purchasing Manager Index (PMI) and PMI new orders improves the nowcasting performance. [103] employ Mixed Data Sampling (MIDAS) framework for the growth rate of GDP in Turkey. Their results show that incorporating daily financial data into the analysis provides better forecasting performance. In order to specify the correct sample, [104] estimates bridge equations using all the combinations of pre-selected 98 indicators. They run approximately four millions of bridge equations to identify the best model for backcasting Turkish GDP. Hence, this approach necessitates conducting computationally demanding and time-consuming efforts. And every time you add a new variable to model, computational cost increases rapidly.

In this chapter, I employ a Dynamic Factor Model (DFM) to forecast Turkish GDP on various samples based on different selection criteria as the literature suggests it. I propose a new sample selection criteria using sparse principal component analysis (SPCA) to construct the best sample data. Also, I empirically compare the performance of my selection method with the widely used selection procedures in the literature. The forecasting exercise performed in a pseudo- real time setting. In particular, I consider the factor model of [105] which is widely used in the recent literature. Moreover, my sample covers the financial crisis of 2008 which may be of great interest to policymakers who would like to assess the forecasting performance of factor models in an episode of high economic stress which makes exercise more challenging.

My results show that the now-casting performance of our dataset which is selected based on SPCA method has the highest level of forecast accuracy for Turkish GDP growth. It delivers the lowest root-mean-square error (RMSE) estimates at all forecast horizons. Moreover, I find that dynamic factor models produce more accurate forecasts compared to the benchmark models. In particular, as more data related to the current quarter becomes available, the forecast accuracy of dynamic factor model increases monotonically with the incorporation of the latest information. Also, modeling of serial correlation of the idiosyncratic component explicitly as an AR(1) process improves the forecasting accuracy of dynamic factor model compared to the one with the assumption of no serial correlation in the idiosyncratic component.

The rest of this chapter proceeds as follows. Section 2 overviews the factor model that I consider in my study. Section 3 describes the sample selection methods and data. Section 4 discusses the empirical results. Finally, Section 5 concludes the section.

3.2 Related Literature

In recent years, various macroeconomic variables are being released at a more disaggregated level at their disposal than before.The incorporation of this vast amount of data to create more accurate forecasts while keeping the empirical framework small has increased the need for dynamic factor models. [106] and [107] were among the first to implement the dynamic factor approach to macroeconomic data. The reason behind the popularity of dynamic factor models is that they are designed to handle large databases by mitigating the dimensionality problem that usually emerges in typical regression models.

However, recently, the factor models in economics were used only for a small number of variables ([108], [109], [110], and [111], since the first approach introduced by [106] and [107] were too restrictive, assuming orthogonality on the idiosyncratic components (exact factor model). As the sample size increases, the assumption that the factors and errors are serially and cross-sectionally uncorrelated do not hold well with economic data. Besides, a significant problem with this approach is that when the number of variables became larger, there occur too many parameters needed to be estimated which makes the problem computationally infeasible. Diffusion index model introduced by [57] relaxed some strict assumptions of former models by taking into account the weak serial correlation of the idiosyncratic components (approximate factor model). In order to derive the common factors, [57] [28] use non-parametric static principal component analysis which is easy to compute. However, the main drawback of the model is that principal components are consistent as ([55]). Hence, factors estimation framework is not well suited for short time series observation. To incorporate dynamics in forecasting, [57] employ an autoregressive model to the factors. To exploit this dynamics structure in factor models explicitly, several alternatives to the static factor modeling have been addressed in the recent literature. The first stream proposed by [83],[112] was based on state-space representation of the model in a time domain. More precisely, they propose two approaches to estimate dynamic factor model. The first one is the so-called approach consisting of principal components and Kalman filtering ([83]). The second one is based on the quasi-maximum likelihood estimation using Expected - Maximization (EM) algorithm([112]). The second stream developed by [58], [113] is based on the spectral domain is also called the generalized dynamic factor model.[2]

3.3 Econometric Methodology

This section briefly explains factor model that is used in my GDP nowcasting framework. I consider the widely used dynamic factor model of [91]. I produce a forecast for quarterly GDP growth rate (year on year), denoted as y_t.

I consider a panel of observable economic variables $X_{i,t}$ where i indicates the cross-section unit $i = 1, ..., N$ and t denotes the time index $i = 1, ..., T$. Each variable in the dataset can be decomposed into a common part and idiosyncratic part, where the common components capture comovement in the data and are driven by a small number of shocks.

Banbura and Modugno model can be described as:

$$X_t = \Lambda F_t + \xi_t, \qquad \xi_t \sim N(0, \Sigma_e), \tag{10}$$

where F_t is an $r \times 1$ vector of unobserved common factors that reflect most of the co-movement in the variables, Λ is a corresponding $N \times r$ factor loading matrix and idiosyncratic disturbances ξ_t has a diagonal covariance matrix Σ_e.

It is assumed that the common factors F_t follow a stationary VAR(p) process driven by the common shocks $N \sim (0, I_r)$ and Ψ_i is $r \times r$ matrices of autoregressive coefficients.

$$F_t = \sum_{i=1}^{k} \Psi_i F_{t-i} + u_t, \qquad u_t \sim N(0, I_r), \tag{11}$$

The idiosyncratic component follows an AR(1) process. The common shocks u_t

[2]See [114], [99], [115] for more detailed literature review.

and the idiosyncratic shocks ϵ_t are assumed to be serially independent and independent of each other over time.

$$\xi_t = \rho\xi_{t-1} + \epsilon_t, \qquad\qquad \epsilon_t \sim N(0, \sigma^2), \tag{12}$$

To produce forecasts of GDP growth, I incorporate the model by combining the monthly factor models in equations (2.1)-(2.3) with a forecast equation for unobserved yearly growth rate of GDP.

$$y_t = \mu + \beta' F_t + \varepsilon_t, \qquad\qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{13}$$

The estimation procedure is quasi-maximum likelihood. Following the Banbura and Modugno (BM) (2014) [105], parameters are estimated by maximum likelihood method using the Expected Maximization (EM) algorithm which can be implemented in the presence of arbitrary patterns of missing data. The main idea of the algorithm is to write the likelihood as if there were no missing data and to iterate between two steps. In the first step, the expectation of the log-likelihood conditional on observed data is computed using the estimates from the previous iteration. In the second step, the parameters are re-estimated by maximizing the expected likelihood with respect to parameters set [3]. To obtain the initial values for the algorithm, firstly, I fill the missing data using the probabilistic PCA method, and then I estimate the factors by using principal components analysis on the complete dataset. Lastly, the remaining parameters are estimated by OLS on the estimated factors.

One of the advantages of BM model is that modeling serial correlation of the idiosyncratic component explicitly can mitigate the misspecification problem that should be handled carefully in small samples. Also, explicit modeling of the idiosyncratic component can be useful to forecast variables with strong non-common dynamics ([91]).

[3] For technical details on the EM iterations and parameter estimates, see [105]

3.4 Sample Selection Methods

In literature, there is no consensus on how we select indicators that are informative for forecasting. Although, factor models can be estimated on large datasets when constructing a dataset one should pay more attention to not including noisy variables. Having too many variables in the dataset is likely to make extraction of relevant signals in the factor model framework more difficult. Moreover, the more the number of variables does not necessarily improve the quality of the forecast. The question of whether there is any gain from selecting a limited number of variables in terms of forecasting performance is critical for practitioners especially. For example, [114] show that a reasonable cross sectional size performs equally well in forecasting in comparison to databases with a very large amount of data. Also, [91] state that factors extracted from large databases is a bit less accurate in forecasting compared to small and medium size specifications. To minimize the impact of uninformative predictors in forecasting using the factor model framework, the literature suggests three possible ways to select informative ones.

1. **Statistical Methods:** These methods can be decomposed into two parts: *Hard Thresholding* and *Soft Thresholding*. Variable selection in the method of hard thresholding is based on some systemic pre-test procedure which leads to the decision of whether a predictor is informative or not. Variables that are above the particular threshold level are included, and variables below it are dropped. [116] shrink full sample series via a simple correlation method. In the first step, all variables with a correlation coefficient of less than 0.5 in absolute terms with respect to GDP growth rates were excluded. In the second step, they dropped the series that have a correlation with any other one above 0.9 in absolute terms. The main disadvantage of this selection method is that it exploits the only bivariate relationship between variables and the series to be forecasted, without

regarding the information included in other indicators. On the other hand, soft thresholding procedures which are based on penalized least squares estimation can perform variable selection and shrinkage simultaneously. [98] implement the least absolute shrinkage selection operator algorithm to end up with a dataset of lower dimension when the idiosyncratic errors are cross correlated. Also, [96] show that soft thresholding methods can be used successfully to reduce the size of the large panel economic data. Depending on the form of penalized function, different types of soft thresholding methods can be proposed. In this chapter, I consider the Least absolute shrinkage selection operator (LASSO). [4] As a consequence, correlation method and LASSO conclude 23 and 47 variables respectively.

2. **Subjective Judgement:** To take advantage of the additional information from a large sample, practitioners use all available variables to forecast GDP. The sample size in the literature varies from 80 to more than 400 variables (See Table 21). For Turkey, my dataset includes 117 macroeconomic time series.

3. **Market Analyst:** This method is pioneered by [105] and followed in [94], [100] and in [101]. The key assumption of this approach is that market participants specialize in choosing the key variables in order to assess the condition of a given country. Bloomberg reports a "relevance index" for each variable that is closely followed by market participants. Thus, we can select the variables based on this index. This data set contains 14 variables.

In this chapter, I propose a new sampling technique based on Sparse Component Analysis (SPCA) that is widely used in the application of image processing, machine learning, and biology. On the other hand, in econometrics, sparse PCA is used for

[4] In the appendix, I provide details on how LASSO penalize the least square estimation.

applications of portfolio selection ([117]). More recently, [118] explore forecast performance of a vast variety of model types including SPCA for the prediction of the key macroeconomic variables (GDP, unemployment, interest rates, consumer price index, etc.). They provide strong evidence on the usefulness of SPCA for factor based forecasting. However, economists have yet to explore its usefulness as a variable selection tool for choosing informative predictors for forecasting, to the best of our knowledge. Since my dataset is characterized by a large number of macroeconomic variables, selecting the appropriate amount of information is very crucial for estimating factor models. Therefore, I adopt SPCA method to extract the informative indicators from my large dataset.

Table 21: Sample Selection - Literature Survey

Paper	Country	Data Sample	Sample Selection Method
Stock and Watson (2002)	U.S.	215 variables	Judgemental
Forni et.al (2005)	Euro Area	447 variables	Judgemental
Boivin and Ng (2006)	U.S.	147 variables	Statistical Methods
Bai and Ng (2008)	U.S.	132 variables	Statistical Methods
Schumacher (2010)	Germany	200 variables	Judgemental
Bańbura and Rünstler (2011)	Euro Area	76 variables	Judgemental
Bessec (2013)	France	96 variables	Statistical Methods
Bańbura and Modugno (2014)	Euro Area	Small - 14 variables Medium - 48 variables Large - 101 variables	Judgemental
Luciani and Ricci (2014)	Norway	14 variables	Market Analyst
Modugno et.al (2014)	Brazil	13 variables	Market Analyst
Li and Chen (2014)	U.S	107 variables	Statistical Methods
Luciani et.al (2015)	Indonesia	12 variables	Market Analyst

3.4.0.1 Sparse Principal Component Analysis

Sparse Principal Component Analysis (SPCA) is a form of the classical Principal Component Analysis (PCA) problem. PCA can produce an estimate of the latent factors, called principal components. Specifically, PCA yields orthogonal vectors that capture the maximum variance in data as possible. One of the drawbacks of this method is that each principal component is a non-zero linear combination of all original variables, which makes practical interpretation of factors difficult. On the other hand, SPCA provides sparse principal components by adding sparsity constraints to the standard PCA framework. Hence, it brings better interpretation by placing zero coefficients on various factor loadings coefficients, i.e. each component is a linear combination of a subset of the original variables.

The sparse PCA problem can be formulated as the following maximization problem:

$$
\begin{aligned}
\underset{X}{\text{maximize}} \quad & v^T (X^T X) v, \\
\text{subject to} \quad & \sum_{j=1}^{N} |v_j| \leqslant \psi, \\
& v^T v = 1.
\end{aligned}
$$

where X is the data matrix, v is the principal components with possible zero loadings and ψ is some tuning parameter. Unfortunately, optimisation of the sparse PCA problems are not trivial, since it is a combinatorial problem. In literature, to perform sparse PCA various algorithms are suggested based on a convex semidefinite programming framwork, generalized power method, greedy search and exact methods using branch-bound techniques.

Following the Naikal et al.(2011), I implement the augmented Langrange multiplier method for extracting the sparse principal components[5]. I select the first factor

[5] Augmented Lagrangian Method (ALM) have recently gained increasing attention due to their

that explains the maximum variance, and I choose variables accordingly so that none of them have non-zero coefficients on the first factor. The SPCA sample includes 19 variables. To give more insights into the forces that drive the first factor, I report the estimated factor coefficients in Figure 34.

rapid convergence. See [119] for details.

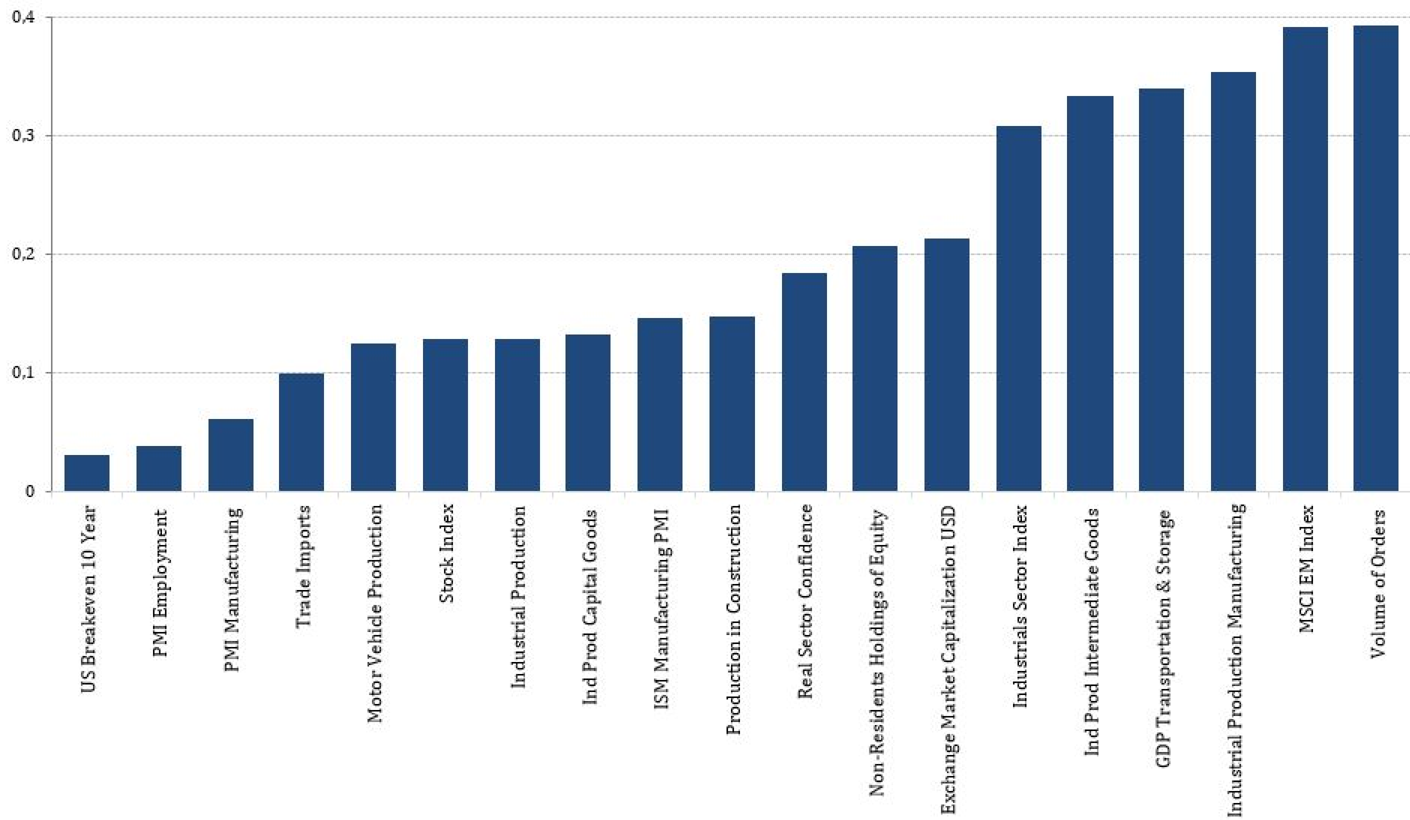


Figure 34: Estimated factor coefficients

3.5 Out of Sample Forecasting Exercise

I use monthly data sets for Turkey. The sample covers the period 2004Q1 - 2015Q2. To compare the performance of factor model on various samples based on different selection criteria, including our own, I perform a pseudo real time out of sample exercise over the period 2008Q1-2015Q2. I call this practice pseudo real time since I use the final data release but take into account the information from each new data releases. For each reference quarter, I estimate a sequence of seven forecasts for GDP growth, starting with the forecast based only on the information available in the first month of the preceding quarter.

For the dynamic factor model (DFM), I followed a similar procedure as in [100] by including two factors (r=2) and two lags (p=2) in the VAR model governing the dynamics of the factors over time. In addition, to satisfy the imposing restriction $q \leq r$, I set q=1.

3.6 Empirical Results

To evaluate the performance of my model and to judge the importance of pre-selecting variables prior to the forecasting, I compare my results with two benchmarks. My first benchmark is a simple autoregressive model of order two on GDP growth. The second one is based on forecasts that extracted from all sample. A ratio smaller than one indicates the improvements in forecast accuracy, related to factor model that incorporate additional information. Table 22 summarizes the relative root mean squared errors (RMSE) of backcasts, nowcasts, and one-quarter forecasts compared to benchmark AR(2) model over the period of 2008Q1 - 2015Q2. The results show that forecast accuracy of factor models deteriorate when the forecast horizon is larger. However, the reduction in RMSE indicates that as more data related to the current quarter becomes available, the forecast accuracy of factor model increases monotonically by incorporating the latest information. Also, it shows that how the model does better than benchmark AR(2) model. Specifically, the factor model starts providing very good forecasting performance at the beginning of the nowcasting period, which is in line with previous researches concluding that factor model is suitable for short term forecasting ([120], [121], [122]). In particular, at the end of the current quarter, the DFM that is estimated from SPCA sample does 32 percent better than the benchmark model, while before GDP is announced it does 47 percent better. This is an important finding since it shows that there is a valuable additional information in the high-frequency data that can be useful to forecast GDP growth.

Table 22: Relative RMSE of GDP forecasts (benchmark AR(2) model). This tables presents the factor model RMSEs of backcasts, nowcasts and one quarter forecasts as ratio to RMSE of the benchmark AR(2) model.

Error Term AR(1) process						
	Month	**BBG**	**All Sample**	**Corr Sample**	**LASSO**	**SPCA**
Forecast	1	1.19	1.32	1.28	1.17	**1.12**
	2	0.98	1.26	1.29	1.20	**0.90**
	3	0.98	1.22	1.29	1.23	**0.95**
Nowcast	1	0.86	1.15	1.10	0.96	**0.85**
	2	0.81	0.99	0.95	0.96	**0.71**
	3	0.76	0.97	0.84	0.96	**0.68**
Backcast	1	0.61	0.73	0.67	0.61	**0.53**

Table 23: Relative RMSE of GDP forecasts (benchmark all sample). This tables presents the factor model RMSEs of backcasts, nowcasts and one quarter forecasts as ratio to RMSE of the forecasts based on all sample.

	Error Term AR(1) process					
	Month	**BBG**	**All Sample**	**Corr Sample**	**LASSO**	**SPCA**
Forecast	1	0.90	1.00	0.97	0.89	**0.85**
	2	0.78	1.00	1.02	0.95	**0.71**
	3	0.80	1.00	1.06	1.01	**0.78**
Nowcast	1	0.75	1.00	0.96	0.83	**0.74**
	2	0.82	1.00	0.96	0.97	**0.72**
	3	0.78	1.00	0.87	0.99	**0.70**
Backcast	1	0.84	1.00	0.92	0.84	**0.73**

Among the selection methods, the SPCA criterion appears to perform best as it attains rank one for all backcast, nowcast and one-quarter ahead forecast periods. All variable selection methods deliver improved forecasts compared to forecasts of all indicators in the data set. Figure 35 , gives further insights into the forecast accuracy results compared to the forecasts based on all sample. [6] Hence, it is crucial to select empirically relevant predictors from a large set of information to mitigate the noise in the data. As suggested by [97], although two researchers use the same model, it is possible to end up with different factor estimates because of the various choices of data. Hence, the selection of correct data set is very important in the nowcasting exercise. And, results confirm that factor model is sensitive to the chosen sample closely.

[6] To avoid complicating the figure too much, I only plot the backcasts.

Table 24: Relative RMSE of GDP forecasts (benchmark AR(2) model). This tables presents the factor model RMSEs of backcasts, nowcasts and one quarter forecasts as ratio to RMSE of the benchmark AR(2) model.

Uncorrelated Error Terms						
	Month	**BBG**	**All Sample**	**Corr Sample**	**LASSO**	**SPCA**
Forecast	1	1.09	1.21	1.28	1.16	**1.08**
	2	1.00	1.19	1.23	1.16	**0.89**
	3	1.01	1.11	1.26	1.19	**0.94**
Nowcast	1	0.88	1.19	1.13	0.98	**0.82**
	2	0.84	1.00	1.01	0.96	**0.76**
	3	0.87	1.01	0.88	0.98	**0.72**
Backcast	1	0.72	0.74	0.68	0.65	**0.60**

In addition, I find improvement area in most of the forecast horizons by projecting the error terms explicitly in the model. In particular, a significant improvement is found in the forecast accuracy of model based on all sample. The main reason is that as more series are included, the possibility of correlated errors will increase. Since asymptotic theories of principal components assume that cross correlation in the errors is not too large, when a sufficient amount of noisy data is added, the average common component will be smaller. As suggested by [97], this creates a situation where more data might not be desirable.

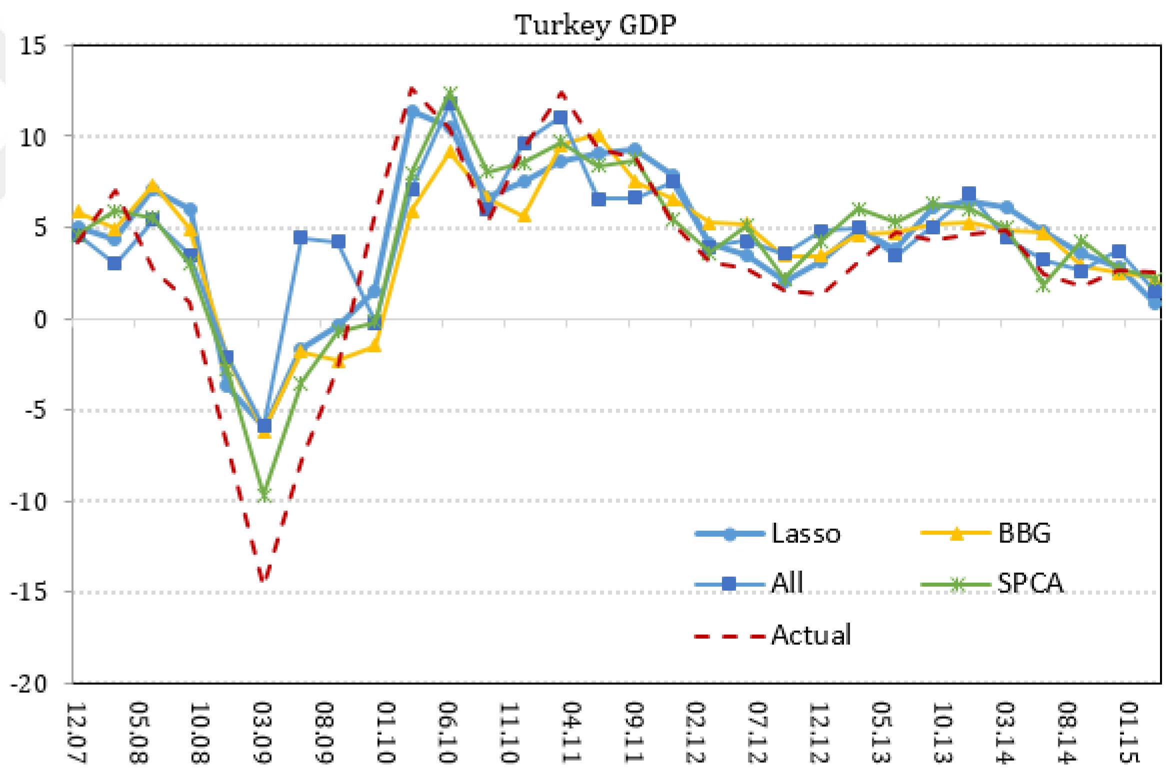


Figure 35: Backcasting of YoY GDP growth

3.7 Conclusion

Factor based prediction has become popular in forecasting literature. Despite the usefulness of summarizing large panel of macroeconomic data, the choice of the sample set from which the factors are extracted is remained partly unaddressed. This chapter provides a comprehensive investigation of this issue by using different sample selection techniques. My results confirm the benefits of incorporating factor-based forecasts and pre-selection of indicators before extracting factors. The focus is on estimating Turkish real GDP in the preceding, current and the next quarter. In order to evaluate the performance of sample selection techniques in nowcasting Turkish GDP, I perform a pseudo real-time forecasting exercise over the period 2008Q1-2015Q2. The results show that factor-based forecasts are sensitive to the sample chosen. I find substantial forecasting gains at all forecast horizons over the benchmark model by estimating the factors using fewer but more informative predictors that are selected using hard and soft thresholding rules.

CHAPTER IV

APPENDIX

4.1 Selection Methods

In this appendix I describe the algorithms used to preselect empirically informative predictors form a set of N potentially relevant indicators in my empirical analysis. I consider a panel of observable economic variables $x_{i,t}$ where i indicates the cross-section unit $i = 1, ..., N$ and t denotes the time index $i = 1, ..., T$. By following the notation of [123], I consider the problem of selecting a subset of X_t is an $T \times N$ matrix for forecasting T dimensional vector Y_t of the yearly growth rate of GDP.

4.1.1 Hard Thresholding

Under hard thresholding, variable selection based on the t-statistic. Variables with t-values above the prescribed threshold level are included, and values below it are dropped. The main disadvantage of this selection method is that it exploits only bivariate relationship between variables and the series to be forecasted, without regarding for the information included in the other indicators. Therefore, it can end up selecting variables that are quite similar and highly correlated with each other.

4.1.2 Soft Thresholding

The traditional hard thresholding displays some discontinuities and may be unstable or more sensitive to small changes in the data because of discreteness of the decision rule. On the other hand, soft thresholding procedures based on penalized least squares estimation can perform variable selection and shrinkage simultaneously. Depending on the form of penalized function, different types of soft thresholding methods can be proposed.

4.1.2.1 LASSO

LASSO (least absolute shrinkage and selection operator) can produce coefficients that are exactly zero and provide a parsimonious model with a few parameters. The shrinkage under LASSO depends on tuning parameter λ that controls the strength of the ℓ_1 penalty.

The LASSO objective function is :

$$\hat{\beta}^{lasso} = \min_{\beta} \quad \|Y - X\beta\|_2 + \lambda \sum_{j=1}^{N} |\beta_j| \tag{14}$$

Since Lasso objective function has absolute value operation, it is not differentiable. As a result, some optimization algorithms must be employed to find the solution of the objective function. For example, an efficient algorithm called the "shooting algorithm" was proposed by [124] that iteratively solves for the LASSO problem in the multi-parameter case. One of the limitation of the LASSO approach is that the number of selected variables is bounded by the number of sample sizes. For example, if N > T, the lasso is able to have at most N non-zero coefficients.

4.2 Economic and Financial Series

Table A lists the names of each economic and financial series, its source and the transformation applied to the series. In the transformation type column, ln denotes logarithm, Δln and $\Delta^2 ln$ denotes the first and second difference of the logarithm, lev denotes the level of the series, and Δlev denotes the first difference of the series.

Brazil/ Number	Name	Source	Transformation
1	Anfavea Brazil Vehicle Production	Anfavea	Δln
2	Anfavea Brazil Vehicle Sales Licensed	Anfavea	Δln
3	Anfavea Brazil Vehicle Exports	Anfavea	Δln
4	Anfavea Brazil Vehicle Sales Licensed Cars	Anfavea	Δln
5	Anfavea Brazil Vehicle Production Passenger	Anfavea	Δln
6	Anfavea Brazil Vehicle Production Trucks	Anfavea	Δln
7	Anfavea Brazil Vehicle Production Buses	Anfavea	Δln
8	Anfavea Brazil Vehicle Production Agricultural	Anfavea	Δln
9	Brazil GDP Market Expectation for end of Current Year Annual Growth	Banco Central do Brasil	Δln
10	Brazil Monthly Economic Activity GDP MOM%	Banco Central do Brasil	lv
11	CNI Brazil Manufacture Industry Capacity Utilization SA	CNI	Δln
12	CNI Brazil Manufacture Industry Capacity Utilization NSA	CNI	Δln
13	CNI Brazil Manufacture Industry Real Sales SA 2006=100	CNI	Δln
14	CNI Brazil Manufacture Industry Employment SA 2006=100	CNI	Δln
15	CNI Brazil Manufacture Industry Working Hours SA 2006=100	CNI	Δln
16	CNI Brazil Industrial Confidence General	CNI	Δln
17	CNI Brazil Consumer Confidence	CNI	Δln
18	Brazil Auto Sales Subtotal	Fenabrave	Δln
19	Brazil Real Industrial Production SA 2002=100	IBGE	Δln

Brazil/ Number	Name	Source	Transformation
20	Brazil Industrial Production Activity Extractive Industry	IBGE	Δln
21	Brazil Industrial Production Activity Manufacturing Industry	IBGE	Δln
22	Brazil Industrial Production Activity Food	IBGE	Δln
23	Brazil Industrial Production Activity Beverage	IBGE	Δln
24	Brazil Industrial Production Activity Tabacco	IBGE	Δln
25	Brazil Industrial Production Activity Textile	IBGE	Δln
26	Brazil Industrial Production Activity Footwear & Leather	IBGE	Δln
27	Brazil Industrial Production Activity Wood	IBGE	Δln
28	Brazil Industrial Production Activity Cellulose & Paper	IBGE	Δln
29	Brazil Industrial Production Capital Goods SA 2002=100	IBGE	Δln
30	Brazil Industrial Production Intermediate Goods SA 2002=100	IBGE	Δln
31	Brazil Industrial Production Consumer Goods SA 2002=100	IBGE	Δln
32	Brazil Industrial Production Durable Goods SA 2002=100	IBGE	Δln
33	Brazil Usually Earned Nominal Total Income	IBGE	Δln
34	Brazil Earned Real Total Income	IBGE	Δln
35	Brazil Retail Sales Volume SA	IBGE	Δln
36	Brazil Retail Sales Volume Furniture and Domestic Appliance SA	IBGE	Δln
37	OECD Brazil Cons. Opin. Confidence Composite & OECD Indicators SA	OECD	Δln
38	OECD Brazil Prod. Manufacturing Total Manufacturing SA 2010=100	OECD	Δln

Brazil/ Number	Name	Source	Transformation
39	OECD Brazil Composite Leading Ind. Total Trend Restored Stck	OECD	$\Delta^2 ln$
40	Brazil GDP YoY 1995=100	IBGE	Δln
41	Brazil Current Account Monthly	Banco Central do Brasil	Δln
42	Brazil BOP Overall Balance	Banco Central do Brasil	Δln
43	Brazil Current Account % of GDP Last 12 Months Accumulated	Banco Central do Brasil	Δln
44	Brazil BOP Financial Account Net	Banco Central do Brasil	Δln
45	Brazil BOP Capital Account Net	Banco Central do Brasil	lev
46	Brazil Foreign Direct Investment Net	Banco Central do Brasil	Δln
47	Brazil Public Net Debt % of GDP	Banco Central do Brasil	lev
48	Brazil Public Primary Budget Result	Banco Central do Brasil	lev
49	Brazil Public Net Debt	Banco Central do Brasil	Δln
50	Brazil General Government Net Debt	Banco Central do Brasil	Δln
51	Brazil BOP Current Account Net	Banco Central do Brasil	Δln
52	Brazil BOP Foreign Direct Investment Net	Banco Central do Brasil	Δln
53	Brazil BOP Portfolio Investment Net	Banco Central do Brasil	lev
54	Brazil BOP Errors and Omissions	Banco Central do Brasil	lev
55	Brazil Total Imports USD	Ministerio da Industria	Δln
56	Brazil Exports USD FOB	Ministerio da Industria	Δln
57	Brazil Trade Balance Weekly Balance	Ministerio da Industria	lev

Brazil/ Number	Name	Source	Transformation
58	Brazil Trade Balance FOB Balance NSA	Ministerio da Industria	lev
59	Brazil Trade Balance FOB Imports NSA	Ministerio da Industria	Δln
60	Brazil Trade Balance FOB Exports	Ministerio da Industria	Δln
61	Brazil Central Government Primary Budget Surplus/Defici	Secretaria do Tesouro Nacional	lev
62	Brazil Central Government Nominal Budget Surplus/Deficit	Secretaria do Tesouro Nacional	lev
63	Brazil Central Government Total Expenditures	Secretaria do Tesouro Nacional	Δln
64	Brazil Government Registered Job Creation	Brazil Labor Ministry	lev
64	Brazil Government Registered Job Creation	Brazil Labor Ministry	lev
65	IBGE Brazil Unemployment Rate Region 30 Days New Methodology	IBGE	lev
66	Brazil Minimum Wage	Ministerio da Fazenda	Δln
67	IMF Brazil Unemployment Rate in Percent per Annu	International Monetary Fund	lev
68	Brazil Average Real Income	IBGE	Δln
68	Brazil Average Real Income	IBGE	Δln
69	Brazil Average Real Income for Public Sector Employees	IBGE	Δln
69	Brazil Average Real Income for Public Sector Employees	IBGE	Δln
70	Brazil Average Real Income for Private Sector Employees	IBGE	Δln
70	Brazil Average Real Income for Private Sector Employees	IBGE	Δln
74	Brazil Average Real Income for Private Sector Full Time Empoy	IBGE	Δln
75	Brazil Average Real Income for Private Sector Non Full Tme Emp	IBGE	Δln

Brazil/ Number	Name	Source	Transformation
76	Brazil Average Real Income for Independent Workers	IBGE	Δln
78	Brazil Unemployment Statistic from 10 Year Total NSA from 6 Major Metropol Areas	IBGE	lev
79	Secovi Sao Paulo Real Estate Units Sale Value	Secovi	Δln
80	Secovi Sao Paulo Real Estate Units Offered	Secovi	Δln
81	Secovi Sao Paulo Real Estate Units Started	Secovi	Δln
82	Secovi Brazil Real Estate Units Average Sale Time Period	Secovi	Δln
83	Secovi Sao Paulo Real Estate Units Sold	Secovi	Δln
84	FGV Brazil Construction Prices INCC-M MoM	Fundacao Getulio Vargas	Δln
85	FGV Brazil Construction Prices INCC-M YoY	Fundacao Getulio Vargas	Δln
86	S&P GSCI Agriculture Index Total Return CME	Standard & Poor's	Δln
87	Brazil CPI IPCA MoM	IBGE	Δln
88	IBGE Brazil CPI Extended National MoM	IBGE	Δln
89	Brazil CPI INPC MoM	IBGE	Δln
90	Brazil CPI IPCA	IBGE	$\Delta^2 ln$
91	S&P GSCI Precious Metals Index Total Return	Standard & Poor's	Δln
92	S&P GSCI Index Spot CME	Standard & Poor's	Δln
93	FGV Brazil Wholesale Prices IPA-M MoM	Fundacao Getulio Vargas	Δln
94	FGV Brazil Wholesale Prices IPA-DI MoM	Fundacao Getulio Vargas	Δln

Brazil/ Number	Name	Source	Transformation
95	Brazil Financial System Loans	Banco Central do Brasil	Δln
96	Brazil Financial Private System Loans	Banco Central do Brasil	Δln
97	Brazil Monetary Base	Banco Central do Brasil	Δln
98	Brazil Money Supply M1 Brazil M1	Banco Central do Brasil	Δln
99	Brazil Money Supply M2 Brazil M2	Banco Central do Brasil	Δln
100	Brazil Money Supply M3 Brazil M3	Banco Central do Brasil	Δln
101	Brazil Money Supply M4 Brazil M4	Banco Central do Brasil	Δln
102	Personal more than 90 days late	Banco Central do Brasil	Δln
103	Brazil Financial System Loans to Housing	Banco Central do Brasil	$\Delta^2 ln$
104	Brazil Financial System Loans to Personal	Banco Central do Brasil	Δln
105	Brazil Intl Daily Reserves	Banco Central do Brasil	Δln
106	USDBRL Spot Exchange Rate - Price of 1 USD in BRL	Banco Central do Brasil	Δln
107	Spread 3M	Banco Central do Brasil	Δln
108	Spread 6M	Banco Central do Brasil	Δln
109	Spread 1 Y	Banco Central do Brasil	Δln
110	Spread 3Y	Banco Central do Brasil	Δln
111	Spread 5Y	Banco Central do Brasil	Δln
112	Brazil Selic Target Rate	Banco Central do Brasil	Δln
113	Brazil Long Term Interest Rate TJLP	Banco Nacional Desenvolvimento	lev

Brazil/ Number	Name	Source	Transformation
114	Andima Brazil Govt Bond Fixed Rate 1 Year	Anbima	Δln
115	Andima Brazil Govt Bond Fixed Rate 2 Years	Anbima	Δln
116	Brazil Government Generic Bond 5 Year	Bloomberg Indices	Δln
117	Brazil Government Generic Bond 10 Year	Bloomberg Indices	Δln
118	Brazil Government Generic Bond 10 Year USD	Bloomberg Indices	Δln
119	Bloomberg Brazil Exchange Market Capitalization USD	Bloomberg Indices	Δln
120	Ibovespa Brasil Sao Paulo Stock Exchange Index	BOVESPA	Δln
121	Brazil Financial Index	BOVESPA	Δln
122	Bovespa Volume Brazil Settlement	Sao Paulo Stock Exchange	Δln

Mexico / Number	Name	Source	Transformation
1	Mexico GDP Total YoY NSA 2008=100	INEGI	Δln
2	Mexico Supply & Demand Private Consumption YoY	INEGI	Δln
3	Mexico Supply & Demand Public Consumption YoY	INEGI	Δln
4	Mexico Economic Indicator Monthly Change	INEGI	lev
5	Mexico Indicator of Economic Activity Index SA	INEGI	Δln
6	Mexico Economic Activity Primary Activities Series Index SA	INEGI	Δln
7	BOM Unit Cost of Labor per Person Employed Manufacturing Industry	Banco de Mexico	Δln
8	Mexico Wholesale/Retail Sale Totl Retl	INEGI	Δln
9	Mexico Wholesale/Retail Sale Totl Whole	INEGI	Δln
10	Mexican Vehicle Sales Auto+truck NSA	AMIA	Δln
11	Mexico Vehicle Production Total Production	AMIA	Δln
12	Industrial Production Total Seasonally Adjusted	INEGI	Δln
13	Industrial Production Mining Seasonally Adjusted	INEGI	Δln
14	Industrial Production Utilities Seasonally Adjusted	INEGI	Δln
15	Industrial Production Construction Seasonally Adjusted	INEGI	Δln
16	Industrial Production Manufacturing Seasonally Adjusted	INEGI	Δln
17	Mexico Manufacturing Index SA	IMEF	Δln
18	Mexico Non Manufacturing Index SA	IMEF	Δln
19	Mexico Manufacturing Aggregate Trend Indicator	INEGI	Δln

Mexico / Number	Name	Source	Transformation
20	MX Consumer Confidence Index SA	INEGI	lev
21	MX Economic Situation of the Household within 12M Compared to the Present SA	INEGI	Δln
22	MX Compared Economic Situation with a Year Ago at Present SA	INEGI	lev
23	Mexico Seasonally Adjusted Leading Indicator	INEGI	Δln
24	Mexico Seasonally Adjusted Coincident Indicator	INEGI	Δln
25	Mexico Capital Investment Construction	INEGI	Δln
26	Mexico Trade Balance Exports Monthly Total USD Million	INEGI	Δln
27	Exports by Sector Petroleum Exports Monthly Total USD Million	INEGI	Δln
28	Exports by Sector Non Petroleum Exports Monthly Total USD Million	INEGI	Δln
29	Mexico Trade Balance Imports Monthly Total USD Million	INEGI	Δln
30	Imports by Sector Consumer Goods Monthly Total USD Million	INEGI	Δln
31	Imports by Sector Intermediate Goods Monthly Total USD Million	INEGI	Δln
32	Imports by Sector Capital Goods Monthly Total USD Million	INEGI	Δln
33	Mexico Nominal Current Account Balance	Banco de Mexico	Δln
34	Banco de Mexico Commercial Balance	Banco de Mexico	Δln
35	BOM Balance of Payments Financial Account	Banco de Mexico	Δln
36	BOM Balance of Payments Financial Account Foreign Direct Investment	Banco de Mexico	Δln
37	BOM Balance of Payments Financial Account Portfolio Investment	Banco de Mexico	Δln

Mexico / Number	Name	Source	Transformation
38	Balance of Payments Errors & Omissions	Banco de Mexico	Δln
39	BOM Public Rev & Expend Budgetary Deficit YTD	Banco de Mexico	Δln
40	BOM Publ Sec Expend Budgetary Expenditures YTD	Banco de Mexico	Δln
41	Mexico Trade Balance Monthly Total USD Million	INEGI	lev
42	Mexico Unemployment Rate SA for Workers 14 and Older ENOE	INEGI	Δln
43	Mexico Manufacturing Employment Laborer or Worker	INEGI	Δln
44	Mexico Manufacturing Employment Office Employees	INEGI	lev
45	Mexico Formal Job Temporary & Permanent Workers Total	INEGI	Δln
46	Mexico Formal Job Temporary & Permanent Workers Manufacturing	INEGI	Δln
47	Mexico Formal Job Temporary & Permanent Workers Construction	INEGI	Δln
48	Mexico Formal Job Temporary & Permanent Workers Retail	INEGI	Δln
49	Mexico Formal Job Temporary & Permanent Workers Transportation & Communication	INEGI	Δln
50	Mexico Formal Job Temporary & Permanent Workers Commercial Services	INEGI	Δln
51	BOM Average Nominal Wages per Person Employed Manufacturing Industry	Banco de Mexico	Δln
52	BOM Average Nominal Wages per Person Employed Commerce	Banco de Mexico	Δln
53	BOM Labor Productivity per Person Employed Manufacturing Industry	Banco de Mexico	Δln
54	BOM Labor Productivity per Person Employed Commerce	Banco de Mexico	Δln
55	BOM Unit Cost of Labor per Person Employed Commerce	Banco de Mexico	Δln

Mexico / Number	Name	Source	Transformation
56	Mexico Manufacturing Index New Orders SA	IMEF	Δln
57	Mexico Non Manufacturing Index New Orders SA	IMEF	Δln
58	Mexico House Price Index YoY	Sociedad Hipotecaria Federal	Δln
59	Mexico Bank Lending Mortgages	Banco de Mexico	Δln
60	Mexico Construction Spending Buildings	INEGI	Δln
61	Mexico Construction Spending Total	INEGI	Δln
62	Mexico House Price Index	Sociedad Hipotecaria Federal	Δln
63	Mexico CPI	INEGI	Δln
64	Mexico Core CPI	INEGI	$\Delta^2 ln$
65	Banco de Mexico CPI Index 2010=100 Food Drinksand Tobacco	INEGI	Δln
66	Banco de Mexico CPI Index 2010=100 Non Food Goods	INEGI	Δln
67	Banco de Mexico CPI Index 2010=100 Services	INEGI	Δln
68	Banco de Mexico CPI Index 2010=100 Agriculture	INEGI	Δln
69	Banco de Mexico CPI Index 2010=100 Energy Rates Auth by Govt	INEGI	Δln
70	Mexico Producer Price Index	INEGI	Δln
71	Mexico Producer Price Index Ex Oil	INEGI	Δln
72	S&P GSCI Index Spot CME	Standard & Poor's	Δln
73	S&P GSCI Agriculture Index Total Return CME	Standard & Poor's	Δln
74	S&P GSCI Precious Metals Index Total Return	Standard & Poor's	Δln

Mexico / Number	Name	Source	Transformation
75	Federal Government Net Domestic Debt in Millions of Mexican Pesos	Secretaria de Hacienda	Δln
76	Mexico Public Sector Net External Debt in Millions of U.S. Dollars	Secretaria de Hacienda	Δln
77	Mexican Money Supply M1-M4 M1 YOY %	Banco de Mexico	Δln
78	Mexican Money Supply M1-M4 M2 YOY %	Banco de Mexico	Δln
79	Mexican Money Supply M1-M4 M3 YOY %	Banco de Mexico	Δln
80	Mexican Money Supply M1-M4 M4 YOY %	Banco de Mexico	Δln
81	Mexican Monetary Base Money Base	Banco de Mexico	Δln
82	MEXICO INTERNATIONAL RESERVE IN US $	Banco de Mexico	Δln
83	Mexico Bank Lending Performing Loans	Banco de Mexico	Δln
84	Mexico Bank Lending Performing Consumer Loans	Banco de Mexico	Δln
85	Mexico Bank Lending Performing Comp	Banco de Mexico	Δln
86	Mexico Bank Lending Performing Loans for Non Bank Financial	Banco de Mexico	Δln
87	BOM Development Banks Total Public Demand Deposits Volume	Banco de Mexico	Δln
88	BOM Development Banks Total Public Time Deposits Volume	Banco de Mexico	Δln
90	USDMXN Spot Exchange Rate - Price of 1 USD in MXN	Bloomberg Indices	Δln
89	Spread 3M	Bloomberg Indices	Δln
91	Spread 6M	Bloomberg Indices	Δln
92	Spread 1 Y	Bloomberg Indices	Δln
93	Spread 3Y	Bloomberg Indices	Δln

Mexico / Number	Name	Source	Transformation
94	Spread 5Y	Bloomberg Indices	Δln
95	Bank of Mexico Official Overnight Rate	Banco de Mexico	lev
96	BOM Government Funding Rate Closing Interest Rate	Banco de Mexico	Δln
97	MXN T-BILL 6 MO	Bloomberg Indices	lev
98	MXN T-BILL 1 YR	Bloomberg Indices	Δln
99	Mexico Generic 2 Year	Bloomberg Indices	Δln
100	Mexico Generic 3 Year	Bloomberg Indices	Δln
101	Mexico Generic 5 Year	Bloomberg Indices	Δln
102	Bloomberg Mexico Exchange Market Capitalization USD	Bloomberg Indices	Δln
103	OECD Mexico Share Prices All Shares Broad Total 2010=100	OECD	Δln
104	Mexican Stock Exchange Mexican Bolsa IPC Index	Mexico Stock Exchange	Δln

S.Africa / Number	Name	Source	Transformation
1	South Africa Retail Sales Total Sales Constant Prices SA 2012=100	Statistics South Africa	Δln
2	South Africa Wholesale Trade Constant 2000 Prices SA	Statistics South Africa	Δln
3	NAAMSA South Africa Total Market Sales Level	NAAMSA	Δln
4	South Africa Manufacturing Production SA 2005=100	Statistics South Africa	Δln
5	South Africa Manufacturing Production SA 2005=100 Food & Beverages	Statistics South Africa	Δln
6	South Africa Manufacturing Production SA 2005=100 Textile Leather Footwear	Statistics South Africa	Δln
7	South Africa Manufacturing Production SA 2005=100 Wood Paper Publish Print	Statistics South Africa	Δln
8	South Africa Manufacturing Production SA 2005=100 Petroleum Chemical Prod	Statistics South Africa	Δln
9	South Africa Manufacturing Production SA 2005=100 Petroleum & Nuclear Fuel	Statistics South Africa	Δln
10	South Africa Manufacturing Production SA 2005=100 Iron Steel Nonferrous Metal	Statistics South Africa	Δln
11	South Africa Mining Sales Total Including Gold SA	Statistics South Africa	Δln
12	South Africa Mining Sales Total Excluding Gold SA	Statistics South Africa	Δln
13	South Africa Mining Sales Building Materials SA	Statistics South Africa	Δln
14	South Africa Mining Sales Other Non Metallic Minerals SA	Statistics South Africa	Δln
15	SACCI South Africa Business Confidence	South African Chamber of Business	lev

S.Africa / Number	Name	Source	Transformation
16	Composite Business Cycle Indicator - Leading Indicator	South African Reserve Bank	Δln
17	Composite Business Cycle Indicator - Coincident Indicator	South African Reserve Bank	Δln
18	South Africa Electricity Production Index Year on Year %	Statistics South Africa	Δln
19	South Africa Real GDP Gross Domestic Expenditure SA	South African Reserve Bank	Δln
20	South Africa Consumer Confidence	Bureau For Economic Research	Δln
21	South Africa Consumer Confidence Economic Position in Next 12m	Bureau For Economic Research	Δln
22	South Africa Consumer Confidence Rating of Present Time To Buy Durables	Bureau For Economic Research	Δln
23	South Africa Household Debt to Disposable Income of Households QoQ	South African Reserve Bank	Δln
24	South Africa Nominal Household Disposable Income SA	South African Reserve Bank	$\Delta^2 ln$
25	South Africa Utilization of Production Capacity	Statistics South Africa	Δln
26	South Africa Trade Balance Incl Oil Arms & Bullion	South African Revenue Service	lev
27	South Africa Trade Balance Exports Incl Oil Arms & Bullion	South African Revenue Service	Δln
28	South Africa Trade Export Other Gd	South African Revenue Service	Δln
29	South Africa Trade Balance Imports Incl Oil Arms & Bullion	South African Revenue Service	Δln
30	South Africa Budget Summary National Budget Balance	South Africa National Treasury	lev
31	South Africa Budget Summary National Expenditures	South Africa National Treasury	Δln
32	South Africa Budget Summary National Revenue	South Africa National Treasury	Δln
33	South Africa Budget Summary Net Borrowing Requirement	South Africa National Treasury	lev
34	South Africa Current Account SA	South African Reserve Bank	Δln

S.Africa / Number	Name	Source	Transformation
35	South Africa Current Account SA - Merchandise Exports Free on Board	South African Reserve Bank	$\Delta^2 ln$
36	South Africa Current Account SA - Net Gold Exports	South African Reserve Bank	Δln
37	South Africa Current Account SA - Less Merchandise Imports	South African Reserve Bank	Δln
38	South Africa Current Account SA - Current Transfers Net Receipts	South African Reserve Bank	Δln
39	South Africa Balance of Payments Capital Transfer Account	South African Reserve Bank	Δln
40	South Africa Balance of Payments Financial Account Net Direct Investment	South African Reserve Bank	Δln
41	South Africa Balance of Payments Financial Account Net Portfolio Investment	South African Reserve Bank	Δln
42	Trade Activity Index Employment	South African Chamber of Business	Δln
43	South Africa Kagiso PMI Employment SA	Kagiso Securities	Δln
44	OECD South Africa Competitiveness Indicator Unit Labour Costs	OECD	lev
45	OECD South Africa Labour Comp. Total Manufacturing Unit Labour Cost 2010=100 SA	OECD	Δln
46	South Africa Number of Employees Nonagricultural Industries	Statistics South Africa	Δln
47	South Africa Unemployment Rate (%)	Bloomberg Indices	Δln
48	IMF South Africa Unemployment Rate as a Percent of Total Labor Force	International Monetary Fund	Δln
49	Trade Activity Index Backlog on Orders	South African Chamber of Business	Δln
50	Trade Expectations Index Backlog on Orders	South African Chamber of Business	Δln
51	Trade Activity Index New Orders	South African Chamber of Business	Δln
52	Trade Expectations Index New Orders	South African Chamber of Business	Δln

S.Africa / Number	Name	Source	Transformation
53	SA Recorded Building Plans Total SA	Statistics South Africa	Δln
54	SA Recorded Building Plans Residentual Buildings SA	Statistics South Africa	Δln
55	SA Recorded Building Plans Non-Residentual Buildings SA	Statistics South Africa	Δln
56	SA Recorded Building Plans Additions and Alterations SA	Statistics South Africa	Δln
57	SA Completed Buildings Recorded Total SA	Statistics South Africa	Δln
58	SA Completed Buildings Recorded Residentual Buildings SA	Statistics South Africa	Δln
59	SA Completed Buildings Recorded Non-Residentual Buildings SA	Statistics South Africa	Δln
60	SA Completed Buildings Recorded Additions and Alterations SA	Statistics South Africa	Δln
61	S&P GSCI Index Spot CME	Standard & Poor's	Δln
62	S&P GSCI Agriculture Index Total Return CME	Standard & Poor's	Δln
63	S&P GSCI Precious Metals Index Total Return	Standard & Poor's	Δln
64	South Africa CPI 2012=100	Statistics South Africa	Δln
65	South Africa CPI YoY (2012=100)	Statistics South Africa	Δln
66	South Africa Kagiso PMI Prices NSA	Kagiso Securities	Δln
67	South Africa GDP Deflator YoY	Statistics South Africa	Δln
68	South Africa Total Liquidations	Statistics South Africa	Δln
69	South Africa Liquidations Cos	Statistics South Africa	Δln
70	South Africa Private Credit Extension	South African Reserve Bank	Δln
71	South Africa Private Credit Extension Investments	South African Reserve Bank	Δln

S.Africa / Number	Name	Source	Transformation
72	South Africa Private Credit Extension Bills Discounted	South African Reserve Bank	Δln
73	South Africa Private Credit Extension Total Loans and Advances	South African Reserve Bank	Δln
74	South Africa Private Credit Extension Installment Sales Credit	South African Reserve Bank	Δln
75	South Africa Private Credit Extension Leasing Finance	South African Reserve Bank	lev
76	South Africa Private Credit Extension Mortgage Advances	South African Reserve Bank	Δln
77	South Africa Private Credit Extension Of Which To Households	South African Reserve Bank	Δln
78	South Africa Money Supply M0	South African Reserve Bank	Δln
79	South Africa Money Supply M1	South African Reserve Bank	Δln
80	South Africa Money Supply M2	South African Reserve Bank	Δln
81	South Africa Money Supply M3	South African Reserve Bank	Δln
82	South Africa Net Open Foreign Currency Position	South African Reserve Bank	Δln
83	South Africa Gold Reserves	South African Reserve Bank	Δln
84	South Africa Private Credit Extension YoY	South African Reserve Bank	lev
85	USDZAR Spot Exchange Rate - Price of 1 USD in ZAR	South African Reserve Bank	Δln
86	Spread 3M	South African Reserve Bank	Δln
87	Spread 6M	South African Reserve Bank	Δln
88	Spread 1 Y	South African Reserve Bank	Δln
89	Spread 3Y	South African Reserve Bank	Δln
90	Spread 5Y	South African Reserve Bank	Δln

S.Africa / Number	Name	Source	Transformation
91	Chicago Board Options Exchange SPX Volatility Index	Chicago Board Options Exchange	Δln
92	US 10 Year 3 month Spread	Bloomberg Indices	Δln
93	Moody's Bond Indices Corporate AAA	Moody's Investors Service	Δln
94	Moody's Bond Indices Corporate BAA	Moody's Investors Service	Δln
95	Federal Funds Target Rate US	Federal Reserve	lev
96	ISM Manufacturing PMI SA	Institute for Supply Managemenent	Δln
97	Republic of South Africa	Bloomberg Indices	Δln
98	South Africa Repo Avg Rate	South African Reserve Bank	Δln
99	South Africa Govt Bonds 2 Year Note Generic Bid Yield	Bloomberg Indices	Δln
100	South Africa Govt Bonds 3 Year Note Generic Bid Yield	Bloomberg Indices	Δln
101	South Africa Govt Bonds 5 Year Note Generic Bid Yield	Bloomberg Indices	Δln
102	South Africa Govt Bonds 10 Year Note Generic Bid Yield	Bloomberg Indices	Δln
103	FTSE/JSE Africa Top40 Tradeable Index	FTSE	Δln
104	FTSE/JSE Africa Financials Index	FTSE	Δln
105	FTSE/JSE Africa Basic Materials Index	FTSE	Δln
106	FTSE/JSE Africa Industrials Index	FTSE	Δln
107	FTSE/JSE Africa Gold Mining Index	FTSE	Δln

Turkey / Number	Name	Source	Transformation
1	Turkey Confidence Index Real Sector	Central Bank of Turkey	Δln
2	Turkey Consumer Confidence	State Institute of Statistics	Δln
3	OECD Turkey Comp Leading Indic Trend Res Stock SA	OECD	$\Delta^2 ln$
4	Turkey Industry Turnover 2010=100	State Institute of Statistics	Δln
5	Turkey Industrial Production 2010=100	State Institute of Statistics	Δln
6	Turkey Industrial Production Manufacturing 2010=100	State Institute of Statistics	Δln
7	Turkey Industrial Production Mining 2010=100	State Institute of Statistics	Δln
8	Turkey Industrial Production Electricity 2010=100	State Institute of Statistics	Δln
9	Turkey Motor Vehicle Industry Production Total	OSD	Δln
10	Turkey Capacity Utilization NSA	Central Bank of Turkey	Δln
11	Turkey Balance of Payments Portfolio Investment Liabilities	Central Bank of Turkey	lev
12	Turkey Balance of Payments Net Errors & Omissions	Central Bank of Turkey	lev
13	Turkey Balance of Payments Direct Investment in Turkey	Central Bank of Turkey	lev
14	Turkey Domestic Debt Position Total	Central Bank of Turkey	Δln
15	Turkey Budget Deficit Primary Balance Before Interest	Republic of Turkey Treasury	lev
16	Turkey Trade Balance	State Institute of Statistics	Δln
17	Turkey Total Exports	State Institute of Statistics	Δln
18	Turkey Total Imports	State Institute of Statistics	Δln
19	Turkey Balance of Payments Current Account	Central Bank of Turkey	Δln

Turkey / Number	Name	Source	Transformation
20	Turkey GDP Constant Prices	State Institute of Statistics	Δln
21	Household Consumption - Total	State Institute of Statistics	lev
22	Household Consumption - Housing	State Institute of Statistics	lev
23	Household Consumption - Durables	State Institute of Statistics	lev
24	Household Consumption - Services	State Institute of Statistics	lev
25	Turkey Labor Statistics Unemployment Rate SA	State Institute of Statistics	Δln
26	Turkey Labor Statistics Employment Rate SA	State Institute of Statistics	Δln
27	Agriculture Forestry Hunting & Fishing	State Institute of Statistics	Δln
28	Mining & Quarrying	State Institute of Statistics	Δln
29	Manufacturing	State Institute of Statistics	Δln
30	Electricity Gas & Water	State Institute of Statistics	Δln
31	Turkey Construction in Thousands	State Institute of Statistics	Δln
32	Wholesale & Retail Trade	State Institute of Statistics	lev
33	Transportation & Communication	State Institute of Statistics	Δln
34	Finance & Insurance	State Institute of Statistics	Δln
35	Community Social & Personal Services	State Institute of Statistics	lev
36	Turkey Unemployment Non-institutional Civilian Population	State Institute of Statistics	$\Delta^2 ln$
37	Turkey Unemployment Labor Force	State Institute of Statistics	Δln
38	Turkey Unemployment Employed	State Institute of Statistics	Δln

Turkey / Number	Name	Source	Transformation
39	Turkey Unemployment Monthly	State Institute of Statistics	Δln
40	Turkey Unemployment Labor Force Participation Rate	State Institute of Statistics	Δln
41	Turkey Unemployment Non-agricultural Unemployment Rate	State Institute of Statistics	Δln
42	Turkey Unemployment Youth Unemployment Rate	State Institute of Statistics	Δln
43	Turkey Unemployment Not in The Labor Force	State Institute of Statistics	Δln
44	OECD Turkey Construction Permits Issued Residential Buildings	OECD	lev
45	Turkey Real Sector Confidence Index Volume of Orders (Current Situation)	Central Bank of Turkey	Δln
46	Turkey Real Sector Confidence Stocks of Finished Goods (Current Situation) SA	Central Bank of Turkey	Δln
47	Turkey Real Sector Confidence Index Export Orders (Next 3 Months) SA	Central Bank of Turkey	Δln
48	Building Permits - State	State Institute of Statistics	Δln
49	Building Permits - Coop	State Institute of Statistics	Δln
50	Building Permits - Private	State Institute of Statistics	lev
51	Building Permits - Total	State Institute of Statistics	lev
52	S&P GSCI Index Spot CME	Standard & Poor's	Δln
53	Turkey PPI	State Institute of Statistics	Δln
54	Turkey PPI Agriculture	State Institute of Statistics	Δln
55	Turkey PPI Industry	State Institute of Statistics	Δln
56	Turkey PPI Crude Petroleum & Natural Gas	State Institute of Statistics	Δln

Turkey / Number	Name	Source	Transformation
57	Turkey PPI Manufacturing	State Institute of Statistics	Δln
58	Turkey PPI Food & Beverages	State Institute of Statistics	Δln
59	Turkey CPI	State Institute of Statistics	Δln
60	Turkey CPI Food & Non Alcoholic Beverages	State Institute of Statistics	Δln
61	Turkey CPI Alcoholic Beverages & Tobacco	State Institute of Statistics	Δln
62	Turkey CPI Housing Water Electricity Gas & Other Fuels	State Institute of Statistics	Δln
63	Turkey CPI Furnishings Household Equipment & Routine House Maintenance	State Institute of Statistics	Δln
64	Turkey CPI Transport	State Institute of Statistics	Δln
65	S&P GSCI Agriculture Index Total Return CME	Standard & Poor's	Δln
66	S&P GSCI Precious Metals Index Total Return	Standard & Poor's	Δln
67	Turkey Consumer Loans Total	Central Bank of Turkey	Δln
68	Deposit Money Banks Loans Private Sector - Housing	Central Bank of Turkey	$\Delta^2 ln$
69	Deposit Money Banks Loans Private Sector - Consumer & Other	Central Bank of Turkey	$\Delta^2 ln$
70	Deposit Money Banks Loans Private Sector	Central Bank of Turkey	Δln
71	Turkey Money Supply M1	Central Bank of Turkey	Δln
72	Turkey New Money Supply M2	Central Bank of Turkey	Δln
73	Turkey New Money Supply M3	Central Bank of Turkey	Δln
74	Turkey Money Supply Time Deposits TRY	Central Bank of Turkey	Δln
75	Turkey Money Supply Sight Deposits FX	Central Bank of Turkey	Δln

Turkey / Number	Name	Source	Transformation
76	Turkey Money Supply Sight Deposits TRY CBRT	Central Bank of Turkey	lev
77	Turkey Money Supply Bank Vaults	Central Bank of Turkey	Δln
78	Turkey Money Supply Sight Deposits TRY	Central Bank of Turkey	Δln
79	Turkish Money Supply Time Deposits FX	Central Bank of Turkey	Δln
80	Turkey Intl Weekly Reserves	Central Bank of Turkey	Δln
81	Turkey Money Supply Repos	Central Bank of Turkey	lev
82	Turkey Money Supply Money Market Funds	Central Bank of Turkey	Δln
83	Weighted Average Interest Rates for Turkish Lira Banks Loans - Commercial	Central Bank of Turkey	Δln
84	USD	Central Bank of Turkey	Δln
85	EUR	Central Bank of Turkey	Δln
86	YEN	Central Bank of Turkey	Δln
87	Swiss franc	Central Bank of Turkey	Δln
88	Implied Vol	Central Bank of Turkey	Δln
89	Risk Revrsal	Central Bank of Turkey	Δln
90	Turkey Real Effective Exchange Rate Broad	BIS	Δln
91	CDS - Country	Bloomberg Indices	Δln
92	CB Rate	Turk Ekonomi Bankasi AS	Δln
93	3 Month	Istanbul Stock Exch Bnd Mrkt	Δln
94	6 month	Istanbul Stock Exch Bnd Mrkt	Δln

Turkey / Number	Name	Source	Transformation
95	1 Year	Istanbul Stock Exch Bnd Mrkt	Δln
96	3 Year	Istanbul Stock Exch Bnd Mrkt	lev
97	5 Year	Istanbul Stock Exch Bnd Mrkt	Δln
98	Weighted Average Interest Rates for Turkish Lira Banks Loans - Cash	Central Bank of Turkey	Δln
99	Weighted Average Interest Rates for Turkish Lira Banks Loans - Vehicles	Central Bank of Turkey	Δln
100	Weighted Average Interest Rates for Turkish Lira Banks Loans - Housing	Central Bank of Turkey	Δln
101	Composite	Istanbul Stock Exchange	Δln
101	Composite	Istanbul Stock Exchange	Δln
102	Banking	Istanbul Stock Exchange	Δln
103	Indutrial	Istanbul Stock Exchange	Δln
104	Utulities	Istanbul Stock Exchange	Δln
106	Turkey Non-Residents Holdings of Equity Stock	Central Bank of Turkey	Δln
107	Turkey Non-Residents Holdings Government Domestic Debt Securities (GDSS) Stock	Central Bank of Turkey	Δln
108	Turkey Eurobond 2 year	Bloomberg Indices	Δln
109	Turkey Eurobond 5 year	Bloomberg Indices	Δln
110	Turkey Eurobond 10 year	Bloomberg Indices	Δln
111	Turkey Eurobond 20 year	Bloomberg Indices	Δln
112	Spread 3M	Bloomberg Indices	Δln

Turkey / Number	Name	Source	Transformation
113	Spread 6M	Bloomberg Indices	Δln
114	Spread 1Y	Bloomberg Indices	Δln
115	Spread 3Y	Bloomberg Indices	Δln
116	Spread 5Y	Bloomberg Indices	Δln

Bibliography

[1] R. Shiller and Y. U. C. F. for Research in Economics, *From Efficient Market Theory to Behavioral Finance.* Cowles Foundation discussion paper: Cowles Foundation for Research in Economics, Cowles Foundation for Research in Economics, Yale University, 2002.

[2] R. C. Merton and Z. Bodie, "The design of financial systems: Towards a synthesis of function and structure," Working Paper 10620, National Bureau of Economic Research, July 2004.

[3] J. H. Cochrane, "Production-based asset pricing and the link between stock returns and economic fluctuations," *Journal of Finance*, vol. 46, pp. 209–37, March 1991.

[4] M. Piazzesi and E. Swanson, "Future prices as risk-adjusted forecasts of monetary policy," *Proceedings*, no. Mar, 2004.

[5] L. Menzly, T. Santos, and P. Veronesi, "Understanding predictability," *Journal of Political Economy*, vol. 112, no. 1, pp. 1–47, 2004.

[6] M. Lettau and S. C. Ludvigson, "Expected returns and expected dividend growth," *Journal of Financial Economics*, vol. 76, pp. 583–626, June 2005.

[7] I. Cooper and R. Priestley, "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, vol. 22, pp. 2801–2833, 2009.

[8] S. Edwards, "The Pricing of Bonds and Bank Loans in International Markets: An Empirical Analysis of Developing Countries' Foreign Borrowing," nber working papers, National Bureau of Economic Research, Inc, Sept. 1986.

[9] W. Cline, K. Barnes, and D. Institute of International Finance (Washington, *Spreads and Risk in Emerging Markets Lending.* Research papers / Institue of International Finance, Institute of International Finance, 1997.

[10] J. Hilscher and Y. Nosbusch, "Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt," *Review of Finance*, vol. 14, no. 2, pp. 235–262, 2010.

[11] M. Uribe and V. Z. Yue, "Country spreads and emerging countries: Who drives whom?," Working Paper 10018, National Bureau of Economic Research, October 2003.

[12] M. Gonzalez-Rozada and E. Yeyati, "Global Factors and Emerging Market Spreads," *Economic Journal*, vol. 118, pp. 1917–1936, November 2008.

[13] S. J. Peiris, "Foreign Participation in Emerging Markets Local Currency Bond Markets ," *IMF Working Paper*, vol. 63, no. WP/10/88, 2010.

[14] D. Duffie, L. H. Pedersen, and K. J. Singleton, "Modeling sovereign yield spreads: A case study of russian debt, forthcoming in journal of finance," 2002.

[15] J. Pan and K. J. Singleton, "Default and recovery implicit in the term structure of sovereign cds spreads. working paper," in *of Sovereign CDS Spreads. Working Paper, MIT Sloan School of Management and Stanford Graduate School of Business*, 2005.

[16] G. Bekaert and C. R. Harvey, "Research in emerging markets finance: looking to the future," *Emerging Markets Review*, vol. 3, pp. 429–448, December 2002.

[17] M. Lee, "Bloomberg fair value market curves," presentation, International Bond Market Conference, 2007.

[18] C. R. Nelson and A. F. Siegel, "Parsimoneous modeling of yield curves for u.s. treasury bills," Working Paper 1594, National Bureau of Economic Research, September 1988.

[19] D. F. Hendry, *Dynamic Econometrics.* No. 9780198283164 in OUP Catalogue, Oxford University Press, 1995.

[20] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[21] T. J. Sargent and C. A. Sims, "Business cycle modeling without pretending to have too much a priori economic theory," working papers, Federal Reserve Bank of Minneapolis, 1977.

[22] J. Geweke, *The Dynamic Factor Analysis of Economic Time Series Models.* SSRI workshop series, Social Systems Research Institute, University of Wisconsin-Madison, 1978.

[23] G. Chamberlain and M. Rothschild, "Arbitrage, factor structure, and mean-variance analysis on large asset markets," Working Paper 996, National Bureau of Economic Research, 1984 1984.

[24] S. C. Ludvigson and S. Ng, "Macro factors in bond risk premia," Working Paper 11703, National Bureau of Economic Research, October 2005.

[25] J. H. Stock and M. W. Watson, "Understanding Changes In International Business Cycle Dynamics," *Journal of the European Economic Association*, vol. 3, pp. 968–1006, 09 2005.

[26] J. Boivin and S. Ng, "Understanding and comparing factor-based forecasts," Working Paper 11285, National Bureau of Economic Research, May 2005.

[27] C. Otrok and C. H. Whiteman, "Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa," *International Economic Review*, vol. 39, pp. 997–1014, November 1998.

[28] J. H. Stock and M. W. Watson, "Macroeconomic forecasting using diffusion indexes," *Journal of Business & Economic Statistics*, vol. 20, no. 2, pp. 147–162, 2002.

[29] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, "The generalized dynamic-factor model: Identification and estimation," *Review of Economics and statistics*, vol. 82, no. 4, pp. 540–554, 2000.

[30] M. Cochrane, John H.and Piazzesi, "Bond risk premia," *American Economic Review*, vol. 95, pp. 138–160, March 2005.

[31] J. Bai and S. Ng, "Tests for skewness, kurtosis, and normality for time series data," *Journal of Business & Economic Statistics*, vol. 23, no. 1, pp. 49–60, 2005.

[32] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting," *Journal of the American Statistical Association*, vol. 100, pp. 830–840, September 2005.

[33] J. Bai and S. Ng, *Large dimensional factor analysis*. Now Publishers Inc, 2008.

[34] J. Bai and S. Ng, "Determining the number of factors in approximate factor models," *Econometrica*, vol. 70, no. 1, pp. 191–221, 2002.

[35] A. Banerjee and M. Marcellino, "Are there any reliable leading indicators for US inflation and GDP growth?," *International Journal of Forecasting*, vol. 22, pp. 137–151, Jan. 2006.

[36] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, "Do financial variables help forecasting inflation and real activity in the euro area?," *Journal of Monetary Economics*, vol. 50, pp. 1243–1255, September 2003.

[37] M. Marcellino, J. H. Stock, and M. W. Watson, "Macroeconomic forecasting in the area: Country specific versus area-wide information," *European Economic Review*, vol. 47, pp. 1–18, February 2003.

[38] J. H. Stock and M. W. Watson, "Understanding changes in international business cycle dynamics," Working Paper 9859, National Bureau of Economic Research, July 2003.

[39] M. J. Artis, A. Banerjee, and M. Marcellino, "Factor forecasts for the uk," *Journal of Forecasting*, vol. 24, no. 4, pp. 279–298, 2005.

[40] S. Eickmeier and C. Ziegler, "How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach," *Journal of Forecasting*, vol. 27, no. 3, pp. 237–265, 2008.

[41] J. B. Taylor, "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, vol. 39, pp. 195–214, December 1993.

[42] B. S. Bernanke and J. Boivin, "Monetary policy in a data-rich environment," *Journal of Monetary Economics*, vol. 50, pp. 525–546, April 2003.

[43] B. S. Bernanke, J. Boivin, and P. Eliasz, "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach," nber working papers, National Bureau of Economic Research, Inc, Jan. 2004.

[44] F. Altissimo, A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin, and G. Veronese, "EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle," cepr discussion papers, C.E.P.R. Discussion Papers, Dec. 2001.

[45] S. B. Aruoba, F. X. Diebold, and C. Scotti, "Real-time measurement of business conditions," *Journal of Business & Economic Statistics*, vol. 27, no. 4, pp. 417–427, 2009.

[46] O. Blanchard, G. Dell'Ariccia, and P. Mauro, "Rethinking Macroeconomic Policy," *Journal of Money, Credit and Banking*, vol. 42, pp. 199–215, 09 2010.

[47] Ö. Akıncı, "Global financial conditions, country spreads and macroeconomic fluctuations in emerging countries," *Journal of International Economics*, vol. 91, no. 2, pp. 358–371, 2013.

[48] J. D. Burger and F. E. Warnock, "Foreign participation in local-currency bond markets," tech. rep., 2004.

[49] H. White, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, vol. 48, pp. 817–38, May 1980.

[50] E. F. Fama and R. R. Bliss, "The Information in Long-Maturity Forward Rates," *American Economic Review*, vol. 77, pp. 680–92, September 1987.

[51] L. P. Hansen and R. J. Hodrick, "Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models," in *Exchange rates and international macroeconomics*, pp. 113–152, University of Chicago Press, 1983.

[52] M. R. Gibbons and W. Ferson, "Testing asset pricing models with changing expectations and an unobservable market portfolio," *Journal of Financial Economics*, vol. 14, no. 2, pp. 217–236, 1985.

[53] W. K. Newey, "Efficient Estimation with Serial Correlation and Lagged Dependent Variables," *Econometric Theory*, vol. 4, pp. 355–356, August 1988.

[54] M. M. Iacoviello, "Housing wealth and consumption," *FRB International Finance Discussion Paper*, no. 1027, 2011.

[55] J. Bai, "Inferential theory for factor models of large dimensions," *Econometrica*, vol. 71, no. 1, pp. 135–171, 2003.

[56] M. Forni, D. Giannone, M. Lippi, and L. Reichlin, "Opening the black box: Structural factor models with large cross sections," *Econometric Theory*, vol. 25, no. 05, pp. 1319–1347, 2009.

[57] J. H. Stock and M. W. Watson, "Forecasting using principal components from a large number of predictors," *Journal of the American statistical association*, vol. 97, no. 460, pp. 1167–1179, 2002.

[58] M. FORNI *et al.*, "The generalized dynamic factor model, one-sided estimation and fore-casting. 2003," 2004.

[59] F. Altissimo, R. Cristadoro, M. Forni, M. Lippi, and G. Veronese, "New eurocoin: Tracking economic growth in real time," *The review of economics and statistics*, vol. 92, no. 4, pp. 1024–1034, 2010.

[60] S. Eickmeier, "Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model," *Journal of Applied Econometrics*, vol. 24, no. 6, pp. 933–959, 2009.

[61] B. S. Bernanke and K. N. Kuttner, "What explains the stock market's reaction to federal reserve policy?," *The Journal of Finance*, vol. 60, no. 3, pp. 1221–1257, 2005.

[62] J. H. Stock and M. W. Watson, "Implications of dynamic factor models for var analysis," tech. rep., National Bureau of Economic Research, 2005.

[63] M. Forni and L. Gambetti, "The dynamic effects of monetary policy: A structural factor model approach," *Journal of Monetary Economics*, vol. 57, no. 2, pp. 203–216, 2010.

[64] M. Zhu, "Housing markets, financial stability and the economy," in *Opening Remarks at the Bundesbank/German Research Foundation/IMF Conference, IMF*, 2014.

[65] K. E. Case and R. J. Shiller, "The efficiency of the market for single-family homes," 1988.

[66] J. Gallin, "The long-run relationship between house prices and rents," *Real Estate Economics*, vol. 36, no. 4, pp. 635–658, 2008.

[67] J. Favilukis, D. Kohn, S. C. Ludvigson, and S. Van Nieuwerburgh, "International capital flows and house prices: Theory and evidence," in *Housing and the Financial Crisis*, pp. 235–299, University of Chicago Press, 2012.

[68] E. L. Glaeser, "A nation of gamblers: Real estate speculation and american history," tech. rep., National Bureau of Economic Research, 2013.

[69] K. E. Case and R. J. Shiller, "Forecasting prices and excess returns in the housing market," *Real Estate Economics*, vol. 18, no. 3, pp. 253–273, 1990.

[70] J. Aizenman and Y. Jinjarak, "Current account patterns and national real estate markets," *Journal of Urban Economics*, vol. 66, no. 2, pp. 75–89, 2009.

[71] K. Adam, P. Kuang, and A. Marcet, "House price booms and the current account," in *NBER Macroeconomics Annual 2011, Volume 26*, pp. 77–122, University of Chicago Press, 2011.

[72] P. Gete, "Housing markets and current account dynamics," *Available at SSRN 1558512*, 2014.

[73] F. Sá, P. Towbin, and T. Wieladek, "Capital inflows, financial structure and housing booms," *Journal of the European Economic Association*, vol. 12, no. 2, pp. 522–546, 2014.

[74] T. Monacelli, "New keynesian models, durable goods, and collateral constraints," *Journal of Monetary Economics*, vol. 56, no. 2, pp. 242–254, 2009.

[75] Z. Liu, P. Wang, and T. Zha, "Land-price dynamics and macroeconomic fluctuations," *Econometrica*, vol. 81, no. 3, pp. 1147–1184, 2013.

[76] V. Bruno and H. S. Shin, "Cross-border banking and global liquidity," *The Review of Economic Studies*, vol. 82, no. 2, pp. 535–564, 2015.

[77] L. Bork and S. V. Møller, "Housing price forecastability: A factor analysis," in *EFA 2012 Copenhagen Meetings Paper*, 2015.

[78] C. André, "A bird's eye view of oecd housing markets," 2010.

[79] H. Hirata, M. A. Kose, C. Otrok, and M. E. Terrones, "Global house price fluctuations: Synchronization and determinants," tech. rep., National Bureau of Economic Research, 2012.

[80] D. Igan and P. Loungani, "Global housing cycles," 2012.

[81] S. Claessens, M. A. Kose, and M. E. Terrones, "How do business and financial cycles interact?," *Journal of International economics*, vol. 87, no. 1, pp. 178–190, 2012.

[82] F. Bräuning and S. J. Koopman, "Forecasting macroeconomic variables using collapsed dynamic factor analysis," *International Journal of Forecasting*, vol. 30, no. 3, pp. 572–584, 2014.

[83] C. Doz, D. Giannone, and L. Reichlin, "A two-step estimator for large approximate dynamic factor models based on kalman filtering," *Journal of Econometrics*, vol. 164, no. 1, pp. 188–205, 2011.

[84] M. P. Clements and A. B. Galvão, "Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the united states," *Journal of Business & Economic Statistics*, vol. 26, no. 4, pp. 546–554, 2008.

[85] W. J. Jansen, X. Jin, and J. de Winter, "Forecasting and nowcasting real gdp: Comparing statistical models and subjective forecasts," 2012.

[86] S. C. Ludvigson and S. Ng, "Macro factors in bond risk premia," *Review of Financial Studies*, vol. 22, no. 12, pp. 5027–5067, 2009.

[87] A. Cesa-Bianchi, L. F. Cespedes, and A. Rebucci, "Global liquidity, house prices, and the macroeconomy: Evidence from advanced and emerging economies," *Journal of Money, Credit and Banking*, vol. 47, no. S1, pp. 301–335, 2015.

[88] O. Akinci and J. Olmstead-Rumsey, "How effective are macroprudential policies? an empirical investigation," 2015.

[89] K. N. Kuttner and I. Shim, "Can non-interest rate policies stabilize housing markets? evidence from a panel of 57 economies," tech. rep., National Bureau of Economic Research, 2013.

[90] A. Krishnamurthy and A. Vissing-Jorgensen, "The effects of quantitative easing on interest rates: channels and implications for policy," tech. rep., National Bureau of Economic Research, 2011.

[91] M. Bańbura, D. Giannone, and L. Reichlin, "Large bayesian vector auto regressions," *Journal of Applied Econometrics*, vol. 25, no. 1, pp. 71–92, 2010.

[92] E. Angelini, M. Bańbura, and G. Rünstler, "Estimating and forecasting the euro area monthly national accounts from a dynamic factor model.," *OECD Journal: Journal of Business Cycle Measurement & Analysis*, vol. 2010, no. 1, 2010.

[93] C. Schumacher, "Forecasting german gdp using alternative factor models based on large datasets," *Journal of Forecasting*, vol. 26, no. 4, pp. 271–302, 2007.

[94] M. Luciani and L. Ricci, "Nowcasting norway," *International Journal of Central Banking*, vol. 10, no. 4, pp. 215–248, 2014.

[95] M. Bessec, "Short-term forecasts of french gdp: A dynamic factor model with targeted predictors," *Journal of Forecasting*, vol. 32, no. 6, pp. 500–511, 2013.

[96] G. Bulligan, M. Marcellino, and F. Venditti, "Forecasting economic activity with targeted predictors," *International Journal of Forecasting*, vol. 31, no. 1, pp. 188–206, 2015.

[97] J. Boivin and S. Ng, "Are more data always better for factor analysis?," *Journal of Econometrics*, vol. 132, no. 1, pp. 169–194, 2006.

[98] J. Bai and S. Ng, "Forecasting economic time series using targeted predictors," *Journal of Econometrics*, vol. 146, no. 2, pp. 304–317, 2008.

[99] M. Luciani, "Forecasting with approximate dynamic factor models: The role of non-pervasive shocks," *International journal of forecasting*, vol. 30, no. 1, pp. 20–29, 2014.

[100] M. Luciani, M. Pundit, A. Ramayandi, and G. Veronese, "Nowcasting indonesia," *Asian Development Bank Economics Working Paper Series*, no. 471, 2015.

[101] D. Bragoli, L. Metelli, and M. Modugno, "The importance of updating: Evidence from a brazilian nowcasting model," 2014.

[102] H. Ç. Akkoyun and M. Günay, "Nowcasting turkish gdp growth," *Central Bank of the Republic of Turkey Working Paper*, no. 12/33, 2012.

[103] B. Sen Dogan and M. Midilic, "Forecasting turkish real gdp growth in a data rich environment," tech. rep., Research and Monetary Policy Department, Central Bank of the Republic of Turkey, 2016.

[104] M. Gunay *et al.*, "" i just ran four million regressions" for backcasting turkish gdp growth," tech. rep., 2015.

[105] M. Bańbura and M. Modugno, "Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data," *Journal of Applied Econometrics*, vol. 29, no. 1, pp. 133–160, 2014.

[106] T. J. Sargent, C. A. Sims, *et al.*, "Business cycle modeling without pretending to have too much a priori economic theory," *New methods in business cycle research*, vol. 1, pp. 145–168, 1977.

[107] J. Geweke, "The dynamic factor analysis of economic time series,[in:] aigner dj, goldberger as (ed.), latent variables in socio–economic models," 1977.

[108] R. Engle and M. Watson, "A one-factor multivariate time series model of metropolitan wage rates," *Journal of the American Statistical Association*, vol. 76, no. 376, pp. 774–781, 1981.

[109] J. H. Stock and M. W. Watson, "New indexes of coincident and leading economic indicators," in *NBER Macroeconomics Annual 1989, Volume 4*, pp. 351–409, MIT press, 1989.

[110] T. J. Sargent, "Two models of measurements and the investment accelerator," *The Journal of Political Economy*, pp. 251–287, 1989.

[111] D. Quah and T. J. Sargent, "A dynamic index model for large cross sections," in *Business cycles, indicators and forecasting*, pp. 285–310, University of Chicago Press, 1993.

[112] C. Doz, D. Giannone, and L. Reichlin, “A quasi–maximum likelihood approach for large, approximate dynamic factor models,” *Review of economics and statistics*, vol. 94, no. 4, pp. 1014–1024, 2012.

[113] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, “othe generalized dynamic factor model: One% sided estimation and forecasting, pjournal of the american statistical association, 100,” $40, p.\ 830, 2005$.

[114] K. Barhoumi, O. Darné, and L. Ferrara, “Dynamic factor models: A review of the literature,” *Available at SSRN 2291459*, 2013.

[115] J. Breitung and S. Eickmeier, “Dynamic factor models,” *Allgemeines Statistisches Archiv*, vol. 90, no. 1, pp. 27–42, 2006.

[116] K. Arnoštová, D. Havrlant, L. Ržička, and P. Tóth, “Short-term forecasting of czech quarterly gdp using monthly indicators,” *Finance a úvěr (Czech Journal of Economics and Finance)*, vol. 61, no. 6, pp. 566–583, 2011.

[117] J. Fan, J. Lv, and L. Qi, “Sparse high dimensional models in economics,” *Annual review of economics*, vol. 3, p. 291, 2011.

[118] H. H. Kim and N. R. Swanson, “Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods,” *Machine Learning, Variable Selection and Shrinkage Methods (February 2016)*, 2016.

[119] N. Naikal, A. Y. Yang, and S. S. Sastry, “Informative feature selection for object recognition via sparse pca,” in *2011 International Conference on Computer Vision*, pp. 818–825, IEEE, 2011.

[120] D. Giannone, L. Reichlin, and D. Small, “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics*, vol. 55, no. 4, pp. 665–676, 2008.

[121] G. Rünstler, K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze, “Short-term forecasting of gdp using large datasets: a pseudo real-time forecast evaluation exercise,” *Journal of forecasting*, vol. 28, no. 7, pp. 595–611, 2009.

[122] M. Bańbura and G. Rünstler, “A look into the factor model black box: publication lags and the role of hard and soft data in forecasting gdp,” *International Journal of Forecasting*, vol. 27, no. 2, pp. 333–346, 2011.

[123] T. Hastie, R. Tibshirani, and J. Friedman, “Unsupervised learning,” in *The elements of statistical learning*, pp. 485–585, Springer, 2009.

[124] W. J. Fu, “Penalized regressions: the bridge versus the lasso,” *Journal of computational and graphical statistics*, vol. 7, no. 3, pp. 397–416, 1998.