



Priscilla Burity

Essays in Econometrics

Tese de Doutorado

Thesis presented to the Post-graduate Program in Economics of the Departamento de Economia do Centro de Ciências Sociais, PUC–Rio as partial fulfillment of the requirements for the degree of Doutor em Economia

Advisor : Prof. Marcelo Cunha Medeiros

Co–Advisor: Prof. Luciano Vereda

Rio de Janeiro
August 2012



Priscilla Burity

Essays in Econometrics

Thesis presented to the Post-graduate Program in Economics of the Departamento de Economia do Centro de Ciências Sociais, PUC–Rio as partial fulfillment of the requirements for the degree of Doutor em Economia. Approved by the following commission:

Prof. Marcelo Cunha Medeiros

Advisor

Departamento de Economia — PUC–Rio

Prof. Luciano Vereda

Co–Advisor

Faculdade de Economia – UFF

Prof. Juliano Junqueira Assunção

Departamento de Economia – PUC-Rio

Prof. Carlos Viana de Carvalho

Departamento de Economia – PUC-Rio

Prof. Luis Henrique Braido

Escola de Pós-Graduação em Economia – FGV

Prof. Caio Ibsen R.de Almeida

Escola de Pós-Graduação em Economia – FGV

Prof. Mônica Herz

Coordinator of the Social Science Center — PUC–Rio

Rio de Janeiro — August 13, 2012

All rights reserved.

Priscilla Burity

The author graduated in Economics from Universidade Federal do Rio de Janeiro - UFRJ in 2004. She obtained the degree of Mestre at Universidade de São Paulo - USP in 2007.

Bibliographic data

Burity, Priscilla

Essays in Econometrics / Priscilla Burity ; advisor: Marcelo Cunha Medeiros; co–advisor: Luciano Vereda. — 2012.

131 f. : il. ; 30 cm

Tese (Doutorado em Economia)-Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2012.

Inclui bibliografia

1. Economia – Teses. 2. Econometria. 3. Estrutura a Termo. 4. Redes Neurais. 5. Crescimento. 6. Convergência. I. Medeiros, Marcelo Cunha. II. Vereda, Luciano. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 330

Acknowledgments

Aos meu orientadores Professores Marcelo Cunha Medeiros e Luciano, pelos estímulos e ensinamentos. Ao Professor Juliano Assunção, pela fundamental contribuição nos artigos.

Aos meus colegas e amigos da PUC-Rio, pelo apoio e pela intensa (e agradável) convivência. Aos meus colegas e amigos de fora da PUC-Rio, por entenderem minhas ausências.

Aos professores que participaram da banca examinadora.

A todos os professores e funcionários do Departamento pelos ensinamentos e pela ajuda.

À Capes, ao CNPq, à FAPERJ, ao Departamento de Economia da PUC-Rio, e à PUC-Rio pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

À minha família ao Gabriel Hartung pelo amor e carinho em todos os momentos.

Abstract

Burity, Priscilla; Medeiros, Marcelo Cunha; Vereda, Luciano (Advisor). **Essays in Econometrics**. Rio de Janeiro, 2012. 131p. PhD Thesis — Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This work consists of four chapters and is divided into two distinct parts. The first is related to models of the term structure estimated for countries whose sovereign bonds are considered defaultable. The second part is about growth and convergence. In the first article of the first part, we used a no-arbitrage term structure model and chose three eurozone countries for this analysis: Spain, Greece and Italy. We noted that the country's debt has played an important role in the recent expansion of spreads, especially for Greece and Italy. For Spain, the recent increase in spreads has been driven primarily by variables related to Germany (Germany's debt is the most important among them), and market stress (represented by a high-yield index). In the second article of the first part, we use a no-arbitrage model to investigate the determinants of the term structure of sovereign spreads and the expectations of the exchange rate in Brazil. Our exercises suggest that factors related to the Brazilian economy had a positive contribution to the 1999-2006 spreads and, in most of the period, factors related to the U.S. economy had a negative contribution. Since late 2007, this scenario has inverted. Furthermore, we present evidence that the improvement in our external solvency indicator along with the improved inflation scenario are the main responsible for the reduction of spreads from the middle of 2007 on. The first article of the second part proposes a semi-parametric approach to control unobserved heterogeneity in linear regression models. We use extreme neurotic networks estimators and the model is illustrated in an application to the convergence of per capita income across Brazilian municipalities. Finally, in the second article of the second part, we explore recent developments on industrial productivity convergence. We find that productivity growth is negatively related to initial productivity, but that countries with worse institutions have lower level of productivity in equilibrium. The rate of productivity growth also has a nonmonotonic relationship with trade openness and education, being faster at the extremes.

Keywords

Econometrics. Term Structure. Neural Networks. Growth. Convergence.

Resumo

Burity, Priscilla; Medeiros, Marcelo Cunha; Vereda, Luciano (Orientador). **Ensaio em Econometria**. Rio de Janeiro, 2012. 131p. Tese de Doutorado — Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho consiste de quatro artigos e está dividida em duas partes distintas. A primeira está relacionado a modelos de estrutura a termo estimados para países cujos títulos soberanos são considerados sujeitos a default. A segunda parte é sobre crescimento e convergência. No primeiro artigo da primeira parte, usamos um modelo de estrutura a termo afim sem arbitragem e escolhemos três países da zona euro para esta análise: Espanha, Grécia e Itália. Notamos que a dívida do próprio país tem desempenhado um papel importante na ampliação recente dos spreads, sobretudo para a Grécia e Itália. Para a Espanha, o recente aumento dos spreads está sendo impulsionado principalmente pelas variáveis relacionadas com a Alemanha (entre as quais a dívida alemã é a mais importante entre eles), e estresse do mercado (representada por um índice alto rendimento). No segundo artigo da primeira parte, usamos um modelo de sem arbitragem para investigar os determinantes da estrutura a termo dos spreads soberanos brasileiros e as expectativas da taxa de câmbio no Brasil. Nossos exercícios indicam que fatores ligados à economia brasileira tiveram uma contribuição positiva para os spreads 1999-2006 e, na maior parte deste período, os fatores ligados à economia dos EUA tiveram uma contribuição negativa. Desde o final de 2007, esse cenário inverteu. Além disso, apresentam-se evidências de que a melhoria na nossa indicador de solvência externa, juntamente com a melhoria no cenário de inflação, é o principal responsável pela redução de spreads a partir do meio de 2007. O primeiro artigo da segunda parte propõe uma abordagem semi-paramétrica para controlar a heterogeneidade não observada em modelos de regressão linear. Utilizamos estimadores extremos de redes neurais e o modelo é ilustrado em uma aplicação para a convergência de renda per capita entre os municípios brasileiros. Finalmente, no segundo artigo da segunda parte, vamos explorar os desenvolvimentos recentes sobre a convergência para a produtividade industrial. Encontramos que o crescimento da produtividade está negativamente relacionado com a produtividade inicial, mas que países com instituições piores têm um nível de equilíbrio de produtividade mais baixo. A taxa de crescimento da produtividade também tem uma relação não monotônica com a abertura comercial e educação, sendo mais rápida nos extremos.

Palavras-chave

Econometria. Estrutura a Termo. Redes Neurais. Crescimento. Convergência.

Contents

1	European Sovereign Yield Spreads in a Gaussian Macro-Finance Term Structure Model Environment	12
1.1	Introduction	12
1.2	Model	13
1.3	Data	16
1.4	Estimation Strategy	18
1.5	Results	19
1.6	Extending the Sample	22
1.7	Robustness	24
1.8	Conclusion	24
2	Interest and Exchange Rates Integrated Model	44
2.1	Introduction	44
2.2	Underlying model and main assumptions	46
2.3	Data	51
2.4	Results	53
2.5	Conclusion	66
3	Geographical Heterogeneity in Regression Models	68
3.1	Introduction	68
3.2	Semi-Parametric Fixed-Effects Regression Model	69
3.3	Simulations	73
3.4	Applications: Economic Growth and Convergence	74
3.5	Conclusion	78
4	Is the Convergence of the Manufacturing Sector Unconditional?	83
4.1	Introduction	83
4.2	Data	86
4.3	The Underlying Specification and a First Look at Data	87
4.4	Estimation Method	95
4.5	Results	98
4.6	Concluding Remarks	102
A	Riccati Difference Equations	110
B	Proof of Proposition 1	111
C	Data: Quarterly to Monthly Frequency	113
D	Identification Hypothesis	114
E	Bootstrap procedure	116
F	Ricatti difference equations	117
G	Proof of equation (2-10)	118

H	Proof of equation (2-16)	120
I	Proofs of Chapter 3	121
J	Countries in the sample	123
K	Estimation of equation (4-10)	128
L	Versions of tables 4.4 and 4.3 in exercises where the convergence coefficient is estimated in an equation with no dummies and with only period dummies.	129

List of Figures

1.1	1 Year Sovereign Bond Yield Spread (%p.a.) - Selected Countries	30
1.2	Central Government Budget Deficit	31
1.3	Deficit vs. 1 Year Bond Yield Spreads - 4 quarters Moving Average	31
1.4	Moody's High Yield (x-axis) and 10-year bond yields, scatter plots	32
1.5	1 year bond yield spread estimated composition	32
1.6	1 year bond yield spread estimated composition - Only Germany related factors	33
1.7	E^N factor loadings of HY , Ge_{Deb} , IP , $Infl$, Def and Deb for the three countries as a function of maturity N .	34
1.8	Impulse response functions for Italy.	35
1.9	Impulse response functions for Spain.	36
1.10	Impulse response functions for Greece.	37
1.11	Path of Θ_t^N for maturities 1, 5 and 10 years for Italy, Spain and Greece	38
1.12	1 Year Sovereign Bond Yield Spread (%p.a.) - Italy, Spain and Greece. Sample: January 1999 to December 2011	38
1.13	1 Year Sovereign Bond Yield Spread (%p.a.) - Italy, Spain and Greece. Sample: January 1999 to December 2011	39
1.14	E^N factor loadings of HY , Ge_{Deb} , IP , $Infl$, Def , Deb for Italy and Spain as a function of maturity N - Sample: January 1999 to December 2011	39
1.15	Deficit (ex-interest rate payments) as a percentage of GDP	40
1.16	Impulse response functions for Italy - Sample: January 1999 to December 2011	41
1.17	Impulse response functions for Spain - Sample: January 1999 to December 2011	42
1.18	Debt to GDP ratio from 2000Q1 to 2010Q3	43
2.1	1-Year Bond-Yields and Sovereign Spreads	45
2.2	Nominal and Real BRL vs USD exchange rate.	48
2.3	U.S. 1-Year Bond-Yield: Observed and Fitted Values	56
2.4	1-Year USD Bond-Yield Sovereign Spreads: Observed and Fitted Values	57
2.5	1-Year BRL Bond-Yield Sovereign Spreads: Observed and Fitted Values	57
2.6	Estimated Path of Θ for 1-year bond	58
2.7	Effective Fed Funds Rate	59
2.8	Estimated Path of Π	59
2.9	Weights E_N on USD and BRL bond-yield spreads over maturities N . Weights are multiplied by 100.	60
2.10	1-year USD bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by HY and Brazilian and U.S. factors over time.	61

2.11	1-year BRL bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by HY and Brazilian and U.S. factors over time.	61
2.12	1-year USD bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by factors over time.	62
2.13	1-year BRL bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by factors over time.	62
2.14	1-year USD bond-yield spreads: Weights E_{12} multiplied by U.S. factors over time.	63
2.15	1-year BRL bond-yield spreads: Weights E_{12} multiplied by U.S. factors over time.	63
2.16	USD bond-yield spreads: Impulse Responses from a one standard deviation shocks.	64
2.17	BRL bond-yield spreads: Impulse responses from a one standard deviation shocks.	65
3.1	Map of the log(Brazilian per capita income growth 1970-2000)	80
3.2	Map of the log(Brazilian per capita income level 1970)	81
3.3	Growth rate from 1970 to 2000 vs. log(1970 per capita income level)	81
3.4	Conditional Growth rate from 1970 to 2000 vs. log(1970 per capita income level)	82
3.5	Map of the Semi-Parametric Geographical Fixed Effects	82
4.1	β_j 's Histograms	93
4.2	β_{jt} and normalized (zero mean, variance one) values of openness (Panel A), executive constraints (Panel B) and years of schooling (Panel C) measured by the model with latitude, longitude, openness, executive constraints, years of schooling and period and industry dummies.	101
4.3	Partial effects and normalized (zero mean, variance one) values of openness (Panel A), executive constraints (Panel B) and years of schooling (Panel C) measured by the model with latitude, longitude, openness, executive constraints, years of schooling and period and industry dummies. The dashed line is the partial effect of variables in the semi-parametric model and the dotted line is the partial effect as measured in the linear model.	101
4.4	Estimated convergence coefficient median (central mark) and 90% confidence interval (box) by country (period mean) as measured by the model with latitude, longitude and years of schooling. Models with period and industry dummies.	102

List of Tables

1.1	Parameter Estimates - Germany	26
1.2	Parameter Estimates - Italy	27
1.3	Parameter Estimates - Spain	28
1.4	Parameter Estimates - Greece	29
1.5	Autocorrelation of yields, spreads and the three first principal components (PC1, PC2 and PC3) of yields (or spreads, for the IGS countries) orthogonalized with respect to the macroeconomics and risk factors	30
1.6	Debt to GDP Ratio from 2000Q1 to 2010Q3 - Granger Causality Tests	43
2.1	Parameter Estimates - Model for the U.S. Term Structure of Yields	54
2.2	Parameter Estimates - Model for the Brazilian Term Structure of Spreads	55
3.1	Simulation Results: K=1 variable and different combinations of number of observations N and complexity ζ .	78
3.2	Simulation Results: K=2 variables and different combinations of number of observations N and complexity ζ .	79
3.3	Simulation Results: K=3 variables and different combinations of number of observations N and complexity ζ .	79
3.4	Convergence Regressions - Brazilian Municipalities.	80
4.1	Pooled Regressions - 10 year growth rates - 1990-2000 to 1997-2007	90
4.2	Pooled Regressions - 10 year growth rates - 1990 to 2007 - weighted (all countries with the same weight)	91
4.3	$\hat{\beta}_{jt}$'s and selected variables in the first year of the decade (t_0). Estimated coefficients are multiplied by 100.	94
4.4	Country-and-year mean partial effects	100
J.1	Observed Countries, years and number of industries. Trade openness and INDSTST4 available data in common.	124
J.2	Observed Countries, years and number of industries. Executive constraint and INDSTST4 available data in common.	125
J.3	Observed Countries, years and number of industries. Years of schooling and INDSTST4 available data in common.	126
J.4	Observed Countries, years and number of industries. Trade openness, executive constraint, years of schooling and INDSTST4 available data in common.	127
L.1	$\hat{\beta}_{jt}$'s from the linear model and selected variables in the first year of the decade (t_0). Estimated coefficients are multiplied by 100.	130
L.2	Country-and-year mean partial effects	131

1 European Sovereign Yield Spreads in a Gaussian Macro-Finance Term Structure Model Environment

1.1 Introduction

The recent economic crisis, triggered by problems in the subprime mortgage market in the USA, has hit economies around the world. Systemic banking fragility associated with the deterioration of the economic outlook and fiscal stimulus packages left behind quite fragile fiscal positions. The banking crisis has turned into a sovereign debt crisis. The details regarding the origins of the euro zone crisis are discussed in the paper by Durré and Smets (2012) in this book. Here, we explore the fact that increases in countries' fiscal deficit and debt levels can bring an increased perception of sovereign risk. In the sovereign bond markets, this movement can cause higher yields on bonds of these countries to finance their debt.

Following Ang and Piazzesi (2003), we use an arbitrage-free affine term structure model to assess how European sovereign yield spreads (measured as the difference between the relevant countries' bond yield and Germany's bond yield) is affected by fiscal variables. The countries we study are Italy, Greece and Spain. We use monthly data from January 1999 to March 2010 for Italy and Spain, and from January 2001 to March 2010 for Greece. We also check how some of the results change when the sample is extended further, with the deepening of the debt crisis.

In the literature on the European sovereign bond market, the papers tackling the role of fiscal variables on bond yield spreads are not conclusive. Focusing on the European bond market, Bernoth, von Hagen and Schuknecht (2004) and Schuknecht, von Hagen and Wolswijk (2008) studied the relationship between fiscal variables and spreads for some selected maturities.¹ The basic equation of the two articles has spreads as dependent variables and fiscal variables and the maturity of bonds (among other things), as explanatory factors, with no restriction of no-arbitrage. The authors find a positive effect from the deficit and debt spreads. In a more general approach, using a sample of 26 countries, Longstaff et al. (2011) found that the excess returns from investing in sovereign credit are largely compensation for bearing global risk, and that there is little or no country-specific credit risk

¹ It is standard in this literature to evaluate, instead of the bond yields themselves, the spreads in relation to the yields on German securities. Schuknecht, von Hagen and Wolswijk (2008) also studied Canadian government bond market.

premium. Using data from USA, Germany, and Italy, Marattin, Paesani and Salotti (2012) show that fiscal fundamentals significantly affect long term interest rates. Our goal is to use the tools developed by the literature on term structure models summarized above to answer the questions traditionally raised by the literature on the European sovereign bonds market.

The no-arbitrage modeling approach we use has advantages in comparison to an unrestricted Vector Autoregressive (VAR) or single equation approaches, like Bernoth, von Hagen and Schuknecht (2004), Schuknecht, von Hagen and Wolswijk (2008), and Georgoutsos and Migiakis (2010). First, it allows us to assess the impact of fiscal shocks to the whole yield curve, not only on observed yields. VAR models have little to say about how yields to maturities not included in the model move. Second, in VAR models the implied movements of yields in relation to each other may not rule out arbitrage opportunities. Thirdly, the no-arbitrage approach allows us to model the relationship between macroeconomic variables and yield spreads. We know of no other work that addresses the relationship between fiscal variables and the term structure of bond yields for countries in the euro zone in an arbitrage-free term structure model environment.

We believe that the three euro-area countries in this analysis - Spain, Greece and Italy - are representative of the group of European countries that has faced debt financing problems more recently. In summary, our main question is: to what extent can their yield spreads can be attributed to economic fundamentals? In particular, we are interested in the contribution of deficit and debt in the expansion of sovereign spreads in the years after the onset of the economic crisis in 2007. Our idea is to distinguish the effects of fiscal shocks from the effects of shocks to other macroeconomic variables and potentially relevant indicators of risk aversion.

The remainder of this paper is organized as follows. In Section 1.2 we present the model. The data are discussed in the Section 1.3. The estimation strategy is presented in Section 1.4. The results are shown in Section 1.5. In section 1.6 we check how some of the results change when the sample is extended to December 2011, with the deepening of the debt crisis. In Section 1.7, we discuss robustness and finally, Section 1.8 concludes.

1.2 Model

Our strategy consists of estimating a term structure model for Germany and then estimating the spreads between the yields to maturity of bonds from Italy, Greece and Spain (the group we will call IGS) and bonds from Germany. These two steps require the previous estimation of the process for the short rate and the dynamics of factors. In this section, we present the term structure models for Germany (subsection 1.2.1) and the IGS countries (subsection 1.2.2). The short rate

and factors equations are also discussed.

1.2.1 Term Structure Model for Non-Defaultable Bonds

The pricing model of securities issued by the German government which, according to our assumptions, are non-defaultable securities, follows from Ang and Piazzesi (2003). The short rate r_t follows an affine function of all state variables, grouped in the vector \mathbf{X}_t of dimension K ; see Duffie and Kan (1996):

$$r_t = \delta_0 + \delta_1' \mathbf{X}_t. \quad (1-1)$$

The composition of \mathbf{X}_t will be discussed below and includes important fiscal variables. We assume that the standardized (zero mean and unit variance) \mathbf{X}_t follows a first order Gaussian VAR.²

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \Sigma \mathbf{u}_t, \quad (1-2)$$

where \mathbf{u}_t is a Gaussian vector of uncorrelated shocks of variance one.

The price at t of a $(N+1)$ -period bond can be written as:

$$V_t^{(N+1)Ge} = \exp(\bar{A}_{N+1} + \bar{B}'_{N+1} \mathbf{X}_t^{Ge}), \quad (1-3)$$

where \bar{A}_{N+1} and \bar{B}'_{N+1} satisfy well-known Riccati difference equations (see Appendix A for a summary).

1.2.2 Term Structure Model for Defaultable Bonds

In this subsection we consider the model for pricing securities issued by the IGS countries which, according to our assumptions, are defaultable securities. Consider a defaultable bond that at t promises to pay V_{t+N} at maturity date $t+N$ and nothing before that. For any period $s = t$, let: (1-1) $h_s \in [0, 1]$ be the conditional probability at s of default between s and $s + 1$; (1-2) φ_s be the recovery value, in units of account, in case of default; (1-3) M_s be the stochastic discount factor (or pricing kernel) in s of the representative buyer. Therefore, the present value of this bond at t , V_t , is given by; see Duffie and Singleton (1999):

$$V_t = h_t E_t (M_{t+1} \varphi_{t+1}) + (1 - h_t) E_t (M_{t+1} V_{t+1}). \quad (1-4)$$

We make the following assumptions.

Assumption 1: In case of default, the recovery value is proportional to the face value of the bond:

$$\varphi_s = (1 - L_s) V_s, \quad (1-5)$$

where L_s denotes the proportional loss at s .

² The VAR order was chosen by the Schwarz information criterion in a range between 0 and 6.

Assumption 1 follows Pan and Singleton (2008) and must be thought of case-specific. Substituting (1-5) into (1-4), we have (Bonomo and Lowenkron [2008]):

$$\begin{aligned} V_t &= h_t E_t [M_{t+1} (1 - L_{t+1}) V_{t+1}] + (1 - h_t) E_t [M_{t+1} V_{t+1}] \\ &= E_t [M_{t+1} (1 - h_t L_{t+1}) V_{t+1}]. \end{aligned}$$

It is common in this literature to allow liquidity to affect pricing.³ We make the simplifying assumption that illiquidity of the security translates into a fractional cost of rate l . Hence, the total discount rate of the security due to default and illiquidity risks is:

$$V_t = E_t [M_{t+1} \Theta_{t+1} V_{t+1}], \quad (1-6)$$

where $\Theta_{t+1} = 1 - h_t L_{t+1} - l_{t+1}$ is a measure of default and liquidity risk. We assume that $\ln \Theta_{t+1}$ is linear in the factors \mathbf{X}_t , i.e., $\ln \Theta_t := \theta_t = \theta_0 + \theta'_1 \mathbf{X}_t = \theta_0 + \theta'_1 \Phi \mathbf{X}_{t-1} + \theta'_1 \Sigma \varepsilon_t$. We also need to impose an additional assumption:

Assumption 2: For every s , $\Theta_s > 0$.

Assumption 2 states that the discount applied on defaultable bonds cannot be larger than 100%. In the absence of liquidity effects, it means that the case where the probability assigned to default is 100% and, in case of default, the loss rate is 100%, never occurs.

Note that the stochastic discount factor used in the pricing model of a defaultable bond is $M_{t+1} \Theta_{t+1}$, which is the stochastic discount factor M_{t+1} of the representative buyer of the non-defaultable bond (bonds issued by the German government) adjusted for the default and liquidity risk Θ_{t+1} . Once Θ_{t+1} carries information on the recovery rate in case of default, probability of default and liquidity effects on the price of the bond, we should think of Θ_{t+1} as maturity-specific. Therefore, we should rewrite equation (1-6) as:

$$V_t^{N+1} = E_t [M_{t+1} \Theta_{t+1}^N V_{t+1}^N]. \quad (1-7)$$

We assume the following simplifying assumption about $\ln \Theta_t^N$:

Assumption 3: $\ln \Theta_t^N$ has three additive components: a constant, a component linear in the maturity and a component linear in the state variables:

$$\ln \Theta_t^N := \theta_t^N = \theta_0^N + \theta'_1 \mathbf{X}_t = \theta_{00} + N \theta_{01} + \theta'_1 \mathbf{X}_t. \quad (1-8)$$

A crucial implication of Assumption 3 is that $\ln \Theta_t^N$ varies (linearly) with the maturity, but the effect of the state variables on $\ln \Theta_t^N$ does not depend on the maturity. Given the model described above, the following proposition determines the prices of defaultable bonds. Thereafter, prices and other variables related to a specific country in the group of the IGS will be indexed by i .

³ See Duffie and Singleton (1999) and Duffie, Pedersen and Singleton (2003)

Proposition 1: If $M_{t+1} \Theta_{t+1}^{(N)i} V_{t+1}^{(N)i}$ follows a lognormal distribution and under Assumptions 1 to 3, the restriction of no arbitrage implies that the price at t of a defaultable $(N + 1)$ -maturity bond (in the case we are evaluating, a bond issued by one of the IGS governments) can be written as $v_t^{(N+1)i} = v_t^{(N+1)Ge} + \bar{D}_{N+1}^i + \bar{\mathbf{E}}_{N+1}^i \mathbf{X}_t^i$, where lowercase letters denote variables in *log*, \mathbf{X}_t^i represents the matrix of factors for all countries and:

$$\bar{D}_{N+1}^i = \theta_0^{(N)i} + \bar{D}_N^i + \left(\theta_1^i + \bar{\mathbf{E}}_N^i \right)' \Sigma^i \left[\mathbf{J}' \left(\Sigma^{Ge'} \bar{\mathbf{B}}_N - \lambda_0 \right) + \frac{1}{2} \Sigma^{i'} \left(\theta_1^i + \bar{\mathbf{E}}_N^i \right) \right] \quad (1-9)$$

$$\bar{\mathbf{E}}_{N+1}^i = \left(\theta_1^i + \bar{\mathbf{E}}_N^i \right)' \left(\Phi^i - \Sigma^i \mathbf{J}' \lambda_1 \mathbf{J} \right). \quad (1-10)$$

Here, \mathbf{J} is a selection matrix such that $\mathbf{X}_t^{Ge} = \mathbf{J} \mathbf{X}_t^i$ and $\mathbf{y}_t^{(N)i}$ is given by $y_t^{(N)i} = -\frac{v_t^{(N+1)i}}{N} = y_t^{(N)Ge} + D_N + \bar{\mathbf{E}}_N^i \mathbf{X}_t^i$. The spreads $s_t^{(N)i} = y_t^{(N)i} - y_t^{(N)Ge}$ can be written as

$$s_t^{(N)i} = D_N + \bar{\mathbf{E}}_N^i \mathbf{X}_t^i, \quad (1-11)$$

where $D_N = -\frac{\bar{D}_N}{N}$ and $\bar{\mathbf{E}}_N = -\frac{\bar{\mathbf{E}}_N}{N}$.

Proof: See Appendix B

Spreads are affine functions of factors. Once we have yields on bonds of different maturities, we estimate the parameters by imposing no-arbitrage restrictions on cross-section estimates given by equations (1-9) and (1-10). For the defaultable bonds model, we need to estimate the parameters $\psi^i = (\Phi^i, \Sigma^i, \theta_{00}^i, \theta_{01}^i, \theta_1^i)$, for $i = \text{Italy (It), Spain (Sp) and Greece (Gr)}$. Note that θ_{00}^i and θ_{01}^i are scalars, θ_1^i is a K^i -dimensional vector, and Φ^i and Σ^i are $K^i \times K^i$ matrices. Hence, with no restrictions imposed, ψ^i can be written as a $(2+K^i+2K^{i2})$ -dimensional vector of parameters.⁴

1.3 Data

The countries we study are Germany, Italy, Greece and Spain. We use monthly data from January 1999 to March 2010 for Italy and Spain, and from January 2001 to March 2010 for Greece.⁵ We use prices of zero-coupon bonds with maturities of 3, 6, 12, 24, 36, 48, 60, 84 and 120 months, extracted from Bloomberg. Factors are divided into macroeconomic variables, a risk indicator and yields-related variables. The macroeconomic variables are: Industrial production (*IP*), inflation (*I*), deficit (excluding interest payments) as a ratio of GDP (*Def*), and debt as a ratio of GDP (*Deb*).⁶ The high yield risk indicator is Moody's Baa Corporate Bond Yield

⁴ Note that, if some variables that are included in the German model are also included in the defaultable bonds model, Ψ^i and Ψ^{Ge} share some of the VAR parameters of the country i (Φ^i and Σ^i). Parameters in Ψ^{Ge} are presented in Table 1.1.

⁵ Greece joined the European Monetary Union in 2001.

⁶ The ideal fiscal variables would be the expected path of deficit and debt within a period ahead (say, twelve months), because fiscal positions usually are partially anticipated. In the absence of this ideal variable, we use variable themselves which is certainly a caveat.

(*HY*). Industrial production and inflation are measured as the annual changes of the respective indices.

The yields-related variables are the three first principal components of yields (for Germany) or spreads (for the IGS countries) estimated from an eigenvalue decomposition of the observed yields (spreads) covariance matrix, orthogonalized with respect to the macroeconomic and risk factors; see Cochrane and Piazzesi (2008). These yields and spreads factors represent variables other than factors considered above. It was shown by Litterman and Scheinkman (1991) and Ang and Piazzesi (2003) that three latent factors are appropriate to capture most salient features of the yield curve.⁷

As a practical matter, these remaining factors have the role of capturing most of the persistence of yields (and spreads, for the IGS countries). As we see in Table 1.5, a large portion of the autocorrelation of yields and spreads is captured by the first three first principal components of yields orthogonalized with respect to the macroeconomic and risk factors.

Sources for the macroeconomic variables are Eurostat (all variables except *HY*) and the Board of Governors of the Federal Reserve System (*HY*). Fiscal Balance, Debt and GDP monthly data are not available. Monthly data for these variables were constructed from their quarterly observations and from monthly data of some coincident indicators, such as energy consumption, unemployment rate, imports and exports. For details, see Appendix C.

1.3.1 A First Look at Data

In the European case, specifically in the cases of Italy, Greece and Spain, Figures 1.1 to 1.3 suggest that there is a relationship between deficit and sovereign yield spreads. Note in Figure 1 that from 1999 to early 2008, the 1-year yield spread between each of the IGS and Germany sovereign bonds was very low, in accordance with the interest-rate parity theory. From the mid 2008 onwards, this spread began to widen. See in Figure 1.2 that the budget deficit of these selected countries has also increased significantly since 2008. Figure 1.3 (first column) shows the four quarter moving average of deficit and one year bond yield spreads of IGS (the last line shows the mean of these variables among the countries). The figure suggests that deficits are closely related to spreads. Spreads, as well as deficit rates, assumed different dynamics from 2008 on but, very importantly, there is no evident reason to believe that the relationship between deficit and spreads has ever changed.

⁷ A number of recent papers (Joslin, Pribsch and Singleton [2010], Joslin, Le and Singleton [2012], Duffee [2011]) have criticized the assumptions underlying the approach of Ang and Piazzesi (2003) and similar term structure models such as this model, which imply that macroeconomic factors are fully explained by the term structure - a fact that is not empirically supported. This paper does not address these particular issues, which might have a bearing on the results.

1.4 Estimation Strategy

We use the following multi-step estimation procedure:

1. Non-Defaultable Bonds estimation

- (a) We estimate the short rate equation (1-1) and the macro dynamics (1-2) by ordinary least squares (OLS), obtaining $\hat{\delta}_0^{Ge}$, $\hat{\delta}_1^{Ge}$, $\hat{\Phi}^{Ge}$ and $\hat{\Sigma}^{Ge}$.
- (b) We estimate the yield equation (A-5) also by OLS imposing restrictions (A-3) and (A-4). In this step, we obtain $\hat{\lambda}_0$ and $\hat{\lambda}_1$. We assume the simplifying assumption that the matrix λ_1 in equation (A-2) is diagonal.⁸

2. Defaultable Bonds

- (a) We estimate the short rate spread equation (1-11) and the analogous macro dynamics (1-2) for the country i by OLS, obtaining $\hat{\delta}_0^i$, $\hat{\delta}_1^i$, $\hat{\Phi}^i$ and $\hat{\Sigma}^i$.
- (b) We estimate the yield equation (1-11) - for $N=1$ - also by OLS using previously estimated parameters $\hat{\Psi}^{Ge}$ and imposing restrictions (1-9) and (1-10). In this step, we obtain $\hat{\theta}_{00}^i$, $\hat{\theta}_{01}^i$ and $\hat{\theta}_1^i$.

In the non-defaultable bonds model, the state variables \mathbf{X}_t^{Ge} are HY_t and all macroeconomic and yields-related variables from Germany, in the order: HY_t , IP_t , CPI_t , Def_t , Deb_t , $PC1_t$, $PC2_t$, $PC3_t$. In the defaultable bonds model, state variables \mathbf{X}_t^i are HY_t and all macroeconomic and yields-related variables from Germany and the country i . \mathbf{X}_t^i includes the first three principal components of spreads for country i . In this model, the variables are ordered the same way. First HY_t , then the German ones, and then the specific country variables (i.e., $K=8$ and $K^i=15$). As some variables included in the Germany non-defaultable model are also included in the defaultable bonds model, (Φ^i, Σ^i) and (Φ^{Ge}, Σ^{Ge}) share some parameters. The way Φ^i and Σ^i are constructed (and the assumptions on them) can be seen in equations (1-12) to (1-13). The estimation of $\hat{\Sigma}^{Ge}$ and $\hat{\Sigma}^i$ requires the estimation of a Structural VAR. It is assumed that Σ is a lower triangular Cholesky matrix such that $Cov(\Sigma \mathbf{u}_t \Sigma') = \Sigma \Sigma'$. See details in Appendix D.

$$\phi_{15 \times 15}^i = \begin{pmatrix} \phi_{8 \times 8}^{Ge} & \mathbf{0}_{8 \times 7} \\ \phi_{7 \times 15}^i & \end{pmatrix} \quad (1-12)$$

⁸ Results are not very sensitive to this restriction on λ_1 . As an exercise, we estimate the model imposing that λ_1 is block-diagonal as in Ang and Piazzesi (2003) and the results are very similar. Technical reasons made Duffee (2011) also impose restrictions on λ_1 . See Duffee (2011) for a discussion.

$$\Sigma_{15 \times 15}^i = \begin{pmatrix} \Sigma_{8 \times 8}^{Ge} & \mathbf{0}_{8 \times 7} \\ \Sigma_{7 \times 15}^{ii} & \end{pmatrix} \quad (1-13)$$

A possible drawback of this approach is the lack of feedback effects from IGS to Germany. We expect that it is partially solved by the inclusion of the Moody's risk indicator in the model for Germany. See scatter-plots in Figure 1.4. The correlation between Moody's High Yield and yields on bonds of longer maturities is large: for 10-year bond yields, it is 0.83 for Italy, 0.79 for Spain and 0.61 for Greece. The Cholesky identification also precludes simultaneous effects. However, as we are considering monthly data we believe this is a minor problem.

1.5 Results

1.5.1 Parameter Estimates

Tables 1.1 to 1.4 show parameter estimates for the non-defaultable (1.1) and defaultable bonds models (Tables 1.2 for Italy, 1.3 for Spain and 1.4 for Greece). In both models, standard errors are calculated through a non-parametric Bootstrap procedure (using 400 repetitions).⁹ The reported standard errors for the non-defaultable bonds model parameters Ψ^{Ge} take into account all the steps of estimation (steps 1.a and 1.b in Section 4). The reported standard errors for the defaultable bonds model parameters Ψ^i take into account the steps of estimation of the parameters of the model (steps 2.a and 2.b in section 4), taking $\widehat{\Psi}^{Ge}$ as constant.¹⁰

Note that in Tables 1.2 to 1.4 the estimated parameters $\widehat{\theta}_{00}^i$, $\widehat{\theta}_{01}^i$ and $\widehat{\theta}_1^i$, related to the measure of default and liquidity risk, seem to have no statistical significance. In this paper, we are going to focus our analysis on the impulse response functions and the estimated factor loadings across maturities. We believe this result does not rule out the existence of default and liquidity premia because the recent regime (in which investors distinguish between German and IGS debts) comprises only around 20% of the total number of observations in the sample, hampering the task of estimating parameters with precision.

1.5.2 What has been driving spreads?

From equation (1-11), we know that the effect of each factor on the yield curve is determined by the weights \mathbf{E}^N that the term structure model assigns to each spread for bonds of maturity N . Figure 1.5 plots the one year bond yield

⁹ See Horowitz (2001) and Davison and Hinkley (1997). Details in Appendix E.

¹⁰ Note that step 2.b depends on previously estimated $\widehat{\Psi}^{Ge}$. The bootstrap procedure used to estimate the reported standard errors for the defaultable bonds model parameters Ψ^i takes $\widehat{\Psi}^{Ge}$ as constant.

spread estimated decomposition, i.e., the weights of factors in the one year bond yield spread (i.e., for $N=12$) multiplied by the current values of the corresponding factor along time. Factors *Ge* are Germany related variables, and *Fiscal* are deficit and debt. The factors *Others* are industrial production, inflation and the principal components of the spreads.¹¹

In the case of Italy, note that the recent rise in spreads has been driven mainly by Italy's debt and variables related to Germany. Market stress (*HY*) also has played an important role. The Germany related variables have been even more important for the widening of Spain's spreads, with smaller influence of the *Others*, *Debt*, *Deficit* and *HY*. Figure 1.6 plots the weights only of Germany-related factors E^N in the one year bond yield spread multiplied by the current values of the corresponding factor along time. It shows that, for both Italy and Spain, among the Germany related variables, the most important ones for the widening of spreads are Germany's fiscal variables. The results are intuitive: markets see Germany as a lender of last resort for the largest economies in the EMU.

For Greece, the rise in spreads has been driven mainly by Greece's own debt (Figure 1.5), with smaller influence of variables other than factors considered (represented by the principal components of spreads), Greek deficit and Germany-related factors.

1.5.3 Factor loadings along the spread curve

As mentioned, the effect of each factor on the yield curve is determined by the weights E^N that the term structure model assigns on each spread of maturity N . Figure 1.7 plots the weights of *HY*, Germany's debt, industrial production, inflation debt and deficit as a function of yield maturity for the three countries.

A common interesting feature is that the weight of the own country's debt is larger around the maturity of 12 months. For Italy and Spain, Germany's debt pays the most important role until the maturity of 112 and 69 months, respectively. In the case of Greece, the own country's debt is the most relevant variable in the determination of spreads. For all countries, market stress (*HY*) is the variable that has the largest weight over longer maturities (from 113 months on for Italy, 70 months on for Spain and 85 months for Greece). The weight on activity is always negative, which is quite intuitive: economic growth is perceived as reducing sovereign risk.

1.5.4 Impulse Responses

Our term structure model allows us to obtain the response of the yields to shocks at all horizons, including maturities omitted in estimation. The impulse

¹¹ We will focus the analysis in the role of *HY* and fiscal variables.

responses (IRs) for all maturities are known analytical functions of the parameters. Figures 1.8 to 1.10 show the IRs of 6, 12 and 120 months for the three countries. These figures show the movements of the yield curve of different maturities (in rows) in response to one standard deviation shocks in *HY*, Germany's debt, the country's own debt and deficit (in columns).¹²

For Italy (see Figure 1.8), the variable that causes the largest impact is market stress (*HY*), followed by Germany's debt and the country's own debt (*Deb*). A shock of one standard deviation to *HY*, for example, causes an initial response of the one year yield spreads of 35% of its standard deviation ($0.023e-2$ of $0.066e-2$).¹³ A shock of one standard deviation to *Ge_Deb* causes an initial response of the 1 year yield spreads of 34% of its standard deviation ($0.022e-2$) and, from the country's debt, 27% ($0.018e-2$).¹⁴ The responses to shocks to *HY*, Germany's debt and the country's own debt get weaker as the maturity increases.

For Spain, Figure 1.9 shows that Germany's debt is the variable that causes the largest impact, followed by market stress (*HY*) and the country's own debt. A shock of one standard deviation to Germany's debt causes an initial response of the one year yield spreads of 55% of its standard deviation ($0.018e-2$ of $0.032e-2$), from *HY*, 40% ($0.013e-2$) and, from the country's own debt, 14% ($0.004e-2$). Finally, yield spreads of Greece respond more strongly to shocks to debt and deficit. A shock of one standard deviation to *Def* causes an initial response of the 1 year yield spreads of 30% of its standard deviation ($0.08e-2$ of $0.28e-2$) and from *Deb*, 20% ($0.06e-2$).

For all countries, the responses of spreads of different maturities to shocks to inflation are not statistical significant at the 5% level. This result can be explained by the fact that these countries are in a monetary union, and country-specific interest rate is not a monetary policy instrument. Only for Greece do the responses of spreads to activity shocks have statistical significance, again suggesting that economic growth (measured here by industrial production growth) is perceived as reducing sovereign risk.

¹² As the reported parameter estimates, reported standard errors for the impulse responses (defaultable bonds model) takes into account the steps of estimation of the parameters of the model (steps 2.a and 2.b in section 4.1), taking $\hat{\varphi}^{Ge}$ as constant.

¹³ Such small numbers are expected: Recall that $s_t^{(N)i} \equiv y_t^{(N)i} - y_t^{(N)Ge}$, where $y_t^{(N)i} = \ln\left(1 + \frac{Y_t^{(N)i}}{100}\right)$. If Italy's 1 year bond yield is, say, 3% per year, $Y_t^{(12)It} = 3$. If Germany's 1 year bond yield is, say, 2.7% per year, $Y_t^{(12)Ge} = 2.7$ and $s_t^{(12)It} \equiv y_t^{(12)It} - y_t^{(12)Ge} = \ln(1.030) - \ln(1.027) \cong 0.003$.

¹⁴ The analysis based on standard deviation must be taken with caution because Figure 1.1 indicates that the spreads' standard deviation changes greatly with time.

1.5.5 Path of Theta

From equation (1-6), we know that Θ_t is a measure of default risk. Keeping l_t constant, the smaller Θ_t , the larger $h_t L_t$, i.e., the larger the probability of default and/or the proportional loss in case of default. We allowed Θ_t to change over time and across maturities. The path of $\hat{\Theta}_t$ for maturities of 1, 5 and 10 years for Italy, Spain and Greece are shown in Figure 1.11.

Over the whole period of analysis (and with absolutely no restrictions on the estimated parameters), $\hat{\Theta}_t$ is around 1 for all maturities for the three countries, indicating that the discount on bond prices because of default and liquidity risks was always of small magnitude. Markets probably expected that the ECB would become a lender of last resort in such an extreme scenario (i.e. if one of its members experienced a severe liquidity constriction), in spite of the no-bailout clause in the EU treaty (Article 104b). Although a clear tendency in Θ_t cannot be seen, the historical lowest levels of Θ_t were reached after 2008 (April and October-2008 for Italy, September-2008 and January-2010 for Spain, and February-2010 and January-2009 for Greece).

An interesting common feature in the path of Θ_t for Italy and Greece is that, the shorter the maturity, the smaller is Θ_t . In the case of Spain, although the difference of Θ_t across maturities is much smaller, the opposite happens: The longer the maturity, the smaller is Θ_t . This difference may be attributed to the different roles played by liquidity and default risk across maturities and across countries.

1.6 Extending the Sample

Our sample ends in March 2010, and from then, the successive plans to restore confidence in the euro area have failed. The market cost of borrowing reached unsustainable levels for many banks and a significant number of governments that share the euro. Figure 1.12 plots the one-year sovereign bond yield spreads for Italy, Spain and Greece. The figure is analogous to Figure 1.1, with data extended to December 2011. The vertical dashed line indicates March 2010, the last month included in our original data. Right after March 2010, the spreads on Greece's sovereign bonds took an almost vertical path - the maximum value, reached in December 2011, was of 123.9% p.a. Figure 1.13 plots the one-year sovereign bond yield spreads only for Italy and Spain, for a better visualization of the data on these countries. Sovereign bond yield spreads of these countries also increased dramatically, reaching a maximum of 5.2% for Spain and 6.4% p.a. for Italy in November 2011. Especially in Greece's case, we do not expect that the model in this paper will explain this phenomenon, as it was driven by a series of political factors, especially the reduced executive decision-making capability in the

European Union and euro area institutional framework, which helps to explain the insufficient policy response.

As an exercise, we run our model for Italy and Spain using the data from January 1999 to December 2011. As we will see, the main results did not change. We present two sets of results: factor loadings in Figure 1.14 and impulse response functions in Figures 1.16 and 1.17.¹⁵

Figure 1.14 is analogous to Figure 1.7. For both Italy and Spain, Germany's debt plays the most important role at the short end of the yield curve. For Italy, Germany's debt has the larger weight on bonds of maturities up to one year and over 69 months. For Spain, Germany's debt has a larger weight on bonds of maturities up to six months. The own country's debt is the most relevant variable in the determination of yield spreads on bonds in the middle end of the curve for Italy and also at the long end for Spain. The negative weights on *HY* and deficit are an odd result. The negative weight on the deficit of both countries is probably due to the fact that plans to restore confidence in the euro promoted some reduction in deficit levels at a moment of high uncertainty. Figure 1.18 shows the path of deficits in the three countries.

The bottom line from the comparison of Figures 1.14 and 1.7 is that the countries' debt seems to have gained a larger weight in the determination of spreads. For all countries, market stress (*HY*) is the variable that has the largest weight at longer maturities (from 113 months for Italy, 69 months for Spain and 73 months for Greece). The weight on activity (measured by industrial production) is always negative, which is quite intuitive: economic growth is perceived as reducing sovereign risk.

Looking at Figures 1.16 and 1.17, we note at least three differences in comparison with the impulse responses in Figures 1.8 and 1.9. The first is the short term responses of yield spreads to shock to *HY*, which now is negative. The second is the magnitude of the impact, which is larger now. This second fact is probably due to the higher levels of spreads observed in the final of the sample. Finally, it seems that the nature of the responses is almost the same - for Italy the variables that cause the largest impact are Germany's debt and the country's own debt (*Deb*), and for Spain, Germany's debt is the variable that causes the largest impact, followed by market stress (*HY*) and the country's own fiscal variables. Therefore, for Italy and Spain, it seems that the dynamics of the variables' interaction is pretty much the same when the sample is extended. But, given the vertical path of yield spread on Greece's sovereign bonds last months, we do not expect that the same is true for this country.

¹⁵ Other results are available upon request.

1.7 Robustness

The importance of Germany's debt in the determination of bond yield spreads of Spain and Italy found so far raises valid concerns about the role of regional contagion. Indeed, in our model, fiscal variables from a certain IGS country enter only in the equations of its own spreads and one may be concerned that the trend in German debt levels is picking up something not in the model. A consequence of that would be that shocks to other variables, say, Greece's debt, could be the real important source of developments in Italy's or Spain's sovereign spreads and a possible correlation between Greece's and Germany's variables would produce the wrong result that Germany's fiscal variables are the ones that are important in the determination of bond yield spreads of Spain and Italy.

In order to address this point, we performed Granger causality tests presented in Table 1.6.¹⁶ We can see that we do not reject the null that other countries' debts Granger cause Germany's debt until a lag of two quarters. It indicates that shocks from other countries' debt eventually impact Germany's debt but it takes a while. In Figure 1.18, we see debt levels of all the four countries under study. It indicates that the recent (post 2008) dynamics of Germany's debt level is quite different from the dynamics of Spain's and Greece's debt, suggesting that the trend in Germany's debt is not picking up movements in those countries' debt. We can infer the same when it comes to deficit ex-interest rate levels in Figure 1.15.

1.8 Conclusion

Following Ang and Piazzesi (2003), we use an arbitrage-free affine term structure model in which sovereign bond yield spreads are used as dependent variables in an equation that includes, among others, fiscal variables. Our main question is in what extent these yield spreads can be attributed to economic fundamentals. In particular, we are interested in the contribution of deficit and debt levels in pushing up sovereign spreads throughout the years after the onset of the economic crisis of 2007. Our idea is to distinguish the effects of fiscal shocks from the effects of shocks to other macroeconomic variables and potentially relevant indicators of risk aversion. We chose three euro-area countries for this analysis: Spain, Greece and Italy.

It is important to stress that the aim of this paper is to understand what drives European yield spreads rather than to promote a particular modelling innovation. Factors in the model follow a first-order VAR in levels and identification of shocks is made via a Cholesky decomposition.¹⁷ We searched for some support in literature

¹⁶ These tests were performed with the original quarterly debt data, available only from the first quarter of 2000.

¹⁷ We estimate an unrestricted VAR in levels assuming that the variables are either cointegrated or

to order the factors, but we are obviously subject to criticism about both the ordering of variables and the use of Cholesky decomposition itself. Another possible drawback is the lack of feedback effects from IGS to Germany. We expect that it is partially solved by the inclusion of the Moody's risk indicator in the model for Germany, which seems to be highly correlated to yields on bonds especially of longer maturities.

As in Marattin, Paesani and Salotti (2012), fundamentals matter. Our results show that the country's own debt has been playing an important role in the recent widening of spreads, especially for Greece and Italy. For Italy, the variable that causes the largest impact is the market stress, followed by Germany's debt and the country's own debt. A shock of one standard deviation to *HY* causes an initial response of the one year yield spreads of 35% of its standard deviation. A shock of one standard deviation to *GeDeb* causes an initial response of the 1 year yield spreads of 34% of its standard deviation and, from the country's debt, 27%. The responses to shocks to *HY*, Germany's debt and the country's own debt get weaker as the maturity increases. Yield spreads of Greece respond more strongly to shocks to debt and deficit. A shock of one standard deviation to *Def* causes an initial response of the 1 year yield spreads of 30% of its standard deviation and from *Deb*, 20%. For Spain, the recent rise in spreads is being driven mainly by variables related to Germany (amongst which German debt is the most important one), and market stress (represented by a high yield index). A shock of one standard deviation to Germany's debt causes an initial response of the one year yield spreads of 55% of its standard deviation, from *HY*, 40% and, from the country's own debt, 14%. For all countries, the responses of spreads of different maturities to shocks to inflation are not statistically significant at the 5% level. This result can be explained by the fact that these countries are in a monetary union, and country-specific interest rate is not a monetary policy instrument. Only for Greece do the responses of spreads to activity shocks have statistical significance, again suggesting that economic growth (measured here by industrial production growth) is perceived as reducing sovereign risk.

Figures and Tables

stationary. This assumption is supported by the shape of the estimated impulse response functions. If the variables in the VAR were non stationary and not cointegrated, we would have seen non convergent impulse responses.

Table 1.1: Parameter Estimates - Germany

VAR of Factors								
$X_t^{GE} = \Phi^{GE} X_{t-1}^{GE} + \Sigma^{GE} \epsilon_t^{GE}$								
Matrix Φ^{GE}								
	HY _{t-1}	IP _{t-1} ^{GE}	CPI _{t-1} ^{GE}	Def _{t-1} ^{GE}	Deb _{t-1} ^{GE}	Princ. Comp1 _{t-1} ^{GE}	Princ. Comp2 _{t-1} ^{GE}	Princ. Comp3 _{t-1} ^{GE}
HY _t	0.923*** (0.044)	-0.015 (0.034)	0.058** (0.029)	-0.022 (0.031)	-0.057 (0.040)	0.004 (0.023)	-0.025 (0.023)	-0.04* (0.023)
IP _t ^{GE}	-0.057 (0.052)	0.952*** (0.040)	-0.032 (0.031)	0.024 (0.034)	0.022 (0.047)	0.091*** (0.028)	-0.075*** (0.024)	0.05* (0.026)
CPI _t ^{GE}	-0.062 (0.062)	0.12** (0.048)	0.838*** (0.040)	0.015 (0.044)	-0.035 (0.055)	0.111*** (0.033)	0.058* (0.030)	0.072** (0.032)
Def _t ^{GE}	-0.559*** (0.107)	-0.353*** (0.084)	-0.23*** (0.067)	0.266*** (0.077)	-0.398*** (0.098)	-0.203*** (0.056)	-0.262*** (0.056)	-0.018 (0.060)
Deb _t ^{GE}	-0.012 (0.017)	-0.053*** (0.014)	-0.005 (0.012)	0.009 (0.012)	0.987*** (0.016)	-0.02** (0.010)	-0.007 (0.010)	-0.018** (0.009)
Princ. Comp1 _t ^{GE}	-0.192** (0.094)	-0.104 (0.072)	-0.177*** (0.060)	-0.328*** (0.066)	-0.188** (0.087)	0.801*** (0.050)	-0.061 (0.048)	-0.156*** (0.051)
Princ. Comp2 _t ^{GE}	-0.170 (0.107)	-0.109 (0.084)	0.009 (0.070)	-0.384*** (0.076)	-0.083 (0.101)	-0.103* (0.058)	0.616*** (0.057)	-0.054 (0.059)
Princ. Comp3 _t ^{GE}	0.009 (0.106)	-0.042 (0.089)	-0.017 (0.066)	0.010 (0.073)	-0.004 (0.098)	0.136** (0.059)	-0.031 (0.057)	0.74*** (0.055)
Matrix Σ^{GE}								
	HY	IP ^{GE}	CPI ^{GE}	Def ^{GE}	Deb ^{GE}	Princ. Comp1 ^{GE}	Princ. Comp2 ^{GE}	Princ. Comp3 ^{GE}
HY	0.269*** (0.044)	-	-	-	-	-	-	-
IP ^{GE}	0.009 (0.028)	0.307*** (0.035)	-	-	-	-	-	-
CPI ^{GE}	0.021 (0.034)	-0.081** (0.040)	0.353*** (0.038)	-	-	-	-	-
Def ^{GE}	-0.191*** (0.051)	-0.024 (0.070)	0.036 (0.071)	0.632*** (0.070)	-	-	-	-
Deb ^{GE}	-0.015* (0.009)	-0.013 (0.009)	0.002 (0.009)	0.020 (0.013)	0.106*** (0.014)	-	-	-
Princ. Comp1 ^{GE}	-	-	-	-	-	0.582*** (0.066)	-	-
Princ. Comp2 ^{GE}	-	-	-	-	-	0.103* (0.054)	0.672*** (0.037)	-
Princ. Comp3 ^{GE}	-	-	-	-	-	-0.041 (0.070)	0.17*** (0.062)	0.632*** (0.053)
Short Rate Regression								
$r_t^{GE} = \delta_0^{GE} + \delta_1^{GE} X_{t-1}^{GE} + u_t^{GE}$								
	Vector δ_t^{GE}							
δ_0^{GE}	HY _t	IP _t ^{GE}	CPI _t ^{GE}	Def _t ^{GE}	Deb _t ^{GE}	Princ. Comp1 _t ^{GE}	Princ. Comp2 _t ^{GE}	Princ. Comp3 _t ^{GE}
2.796*** (0.002)	0.108*** (0.003)	0.394*** (0.003)	0.348*** (0.002)	-0.307*** (0.002)	-0.539*** (0.003)	0.421*** (0.002)	0.226*** (0.002)	0.074*** (0.002)
Prices of Risk λ_0 and λ_1								
Matrix λ_0^{GE}								
$\lambda_1^{GE} = \lambda_0^{GE} + \lambda_1^{GE} X_t^{GE}$	HY _t	IP _t ^{GE}	CPI _t ^{GE}	Def _t ^{GE}	Deb _t ^{GE}	Princ. Comp1 _t ^{GE}	Princ. Comp2 _t ^{GE}	Princ. Comp3 _t ^{GE}
3.869 (6.170)	-11.734 (12.744)	-17.574 (20.227)	-7.686 (12.014)	-0.891 (2.064)	1.755 (6.681)	0.260 (0.749)	18.507 (32.530)	
Diagonal of Matrix λ_1^{GE}								
	HY _t	IP _t ^{GE}	CPI _t ^{GE}	Def _t ^{GE}	Deb _t ^{GE}	Princ. Comp1 _t ^{GE}	Princ. Comp2 _t ^{GE}	Princ. Comp3 _t ^{GE}
	-0.120 (0.209)	0.209 (0.202)	0.505 (0.515)	-0.131 (0.161)	0.359 (0.628)	0.078 (0.081)	0.031 (0.389)	1.626 (0.998)
Standard deviations in parenthesis; *10% significance level, **5% significance level, ***1% significance level								

Table 1.2: Parameter Estimates - Italy

Italy
VAR of Factors
 $X_t^i = \mu^i + \Phi X_{t-1}^i + \Sigma \epsilon_t^i$

	Matrix Φ^i														
	HY_{t-1}^i	IP_{t-1}^i	CPI_{t-1}^i	Def_{t-1}^i	Deb_{t-1}^i	$Princ. Comp1_{t-1}^i$	$Princ. Comp2_{t-1}^i$	$Princ. Comp3_{t-1}^i$	IP_{t-1}^i	IP_{t-1}^i	Def_{t-1}^i	Deb_{t-1}^i	$Princ. Comp1_{t-1}^i$	$Princ. Comp2_{t-1}^i$	$Princ. Comp3_{t-1}^i$
IP_t^i	-0.068 (0.062)	0.833*** (0.094)	-0.107** (0.050)	0.043 (0.045)	-0.052 (0.061)	0.1*** (0.032)	-0.101*** (0.030)	0.037 (0.029)	0.121 (0.093)	0.079 (0.055)	0.000 (0.056)	0.179*** (0.053)	-0.071** (0.027)	0.001 (0.028)	-0.064** (0.026)
CPI_t^i	-0.089** (0.042)	-0.015 (0.065)	0.106*** (0.037)	0.044 (0.032)	-0.063 (0.041)	0.044* (0.022)	-0.001 (0.019)	0.032 (0.020)	0.058 (0.062)	0.911*** (0.041)	-0.033 (0.038)	0.086** (0.036)	0.003 (0.020)	0.010 (0.019)	-0.003 (0.019)
Def_t^i	-0.126 (0.105)	-0.439*** (0.166)	-0.102 (0.084)	0.053 (0.074)	0.375*** (0.094)	-0.146*** (0.049)	-0.109** (0.047)	-0.106** (0.049)	-0.038 (0.150)	0.002 (0.093)	0.015 (0.091)	-0.382*** (0.085)	0.049 (0.046)	-0.117** (0.048)	-0.088* (0.048)
Deb_t^i	0.093 (0.051)	-0.082 (0.088)	-0.035 (0.050)	-0.003 (0.039)	0.023 (0.051)	-0.031 (0.028)	-0.010 (0.026)	-0.007 (0.027)	0.052 (0.084)	-0.081 (0.054)	0.023 (0.052)	0.849*** (0.046)	0.027 (0.026)	-0.009 (0.025)	0.015 (0.025)
$Princ. Comp1_t^i$	0.011 (0.141)	0.086 (0.249)	-0.194 (0.130)	-0.231** (0.107)	0.028 (0.080)	0.037 (0.067)	-0.017 (0.067)	0.125** (0.072)	0.008 (0.139)	0.114 (0.139)	0.143 (0.133)	-0.035 (0.072)	0.54*** (0.067)	-0.012 (0.067)	-0.065 (0.067)
$Princ. Comp2_t^i$	-0.546*** (0.147)	0.405 (0.268)	-0.082 (0.124)	-0.103 (0.112)	-0.220 (0.139)	0.203*** (0.075)	0.023 (0.076)	-0.068 (0.075)	-0.805*** (0.241)	0.291* (0.147)	-0.221 (0.143)	0.219* (0.129)	0.064 (0.072)	0.455*** (0.068)	0.019 (0.072)
$Princ. Comp3_t^i$	-0.301* (0.169)	-0.302 (0.277)	-0.027 (0.158)	-0.111 (0.135)	-0.148 (0.166)	0.101 (0.093)	0.059 (0.084)	-0.038 (0.086)	0.039 (0.262)	0.136 (0.162)	-0.013 (0.170)	0.047 (0.155)	0.011 (0.082)	-0.016 (0.085)	0.285*** (0.078)

Matrix Σ^i

	HY_t^i	IP_t^i	CPI_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_{t-1}^i$	$Princ. Comp2_{t-1}^i$	$Princ. Comp3_{t-1}^i$	IP_t^i	IP_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_t^i$	$Princ. Comp2_t^i$	$Princ. Comp3_t^i$
HY_t^i	-0.002 (0.025)	0.001 (0.027)	-0.003 (0.033)	-0.003 (0.029)	-0.004 (0.077)	-0.001 (0.026)	-0.002 (0.029)	0.003 (0.030)	0.302*** (0.068)	-	-	-	-	-	-
IP_t^i	-0.002 (0.025)	0.001 (0.027)	-0.003 (0.033)	-0.003 (0.029)	-0.004 (0.077)	-0.001 (0.026)	-0.002 (0.029)	0.003 (0.030)	0.302*** (0.068)	-	-	-	-	-	-
CPI_t^i	-0.002 (0.019)	-0.001 (0.019)	-0.002 (0.021)	0.000 (0.021)	-0.067 (0.091)	0.001 (0.018)	-0.001 (0.022)	0.003 (0.021)	0.133 (0.106)	0.133 (0.106)	-	-	-	-	-
Def_t^i	0.002 (0.038)	-0.001 (0.043)	-0.003 (0.050)	0.002 (0.046)	-0.027 (0.043)	0.001 (0.043)	-0.001 (0.050)	0.000 (0.054)	-0.034 (0.046)	-0.002 (0.046)	0.524*** (0.051)	-	-	-	-
Deb_t^i	0.000 (0.022)	0.001 (0.025)	0.002 (0.027)	0.000 (0.028)	-0.008 (0.024)	-0.001 (0.025)	-0.002 (0.026)	0.000 (0.027)	0.024 (0.027)	0.004 (0.022)	0.025 (0.026)	0.28*** (0.035)	-	-	-
$Princ. Comp1_t^i$	-	-	-	-	-	-	-	-	-	-	-	0.804*** (0.093)	-	-	-
$Princ. Comp2_t^i$	-	-	-	-	-	-	-	-	-	-	-	-	0.796*** (0.088)	-	-
$Princ. Comp3_t^i$	-	-	-	-	-	-	-	-	-	-	-	-	-	0.92*** (0.115)	-

Short Rate Regression
 $spread_t^i = \delta_0^i + \delta_1^i X_{t-1}^i + u_t^i$

	HY_t^i	IP_t^i	CPI_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_{t-1}^i$	$Princ. Comp2_{t-1}^i$	$Princ. Comp3_{t-1}^i$	IP_t^i	IP_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_t^i$	$Princ. Comp2_t^i$	$Princ. Comp3_t^i$
δ_0^i	0.067*** (0.002)	0.082*** (0.004)	0.000 (0.006)	0.000 (0.003)	-0.006** (0.003)	0.135*** (0.004)	-0.029*** (0.002)	0.003 (0.002)	-0.032*** (0.002)	-0.022*** (0.006)	0.04*** (0.004)	-0.02*** (0.004)	0.004 (0.003)	0.045*** (0.002)	0.024*** (0.002)

Recovery Intensity
 $\ln(\theta^i) = \theta_{\infty}^i + \theta_{\infty}^i X_{t-1}^i$

	HY_t^i	IP_t^i	CPI_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_{t-1}^i$	$Princ. Comp2_{t-1}^i$	$Princ. Comp3_{t-1}^i$	IP_t^i	IP_t^i	Def_t^i	Deb_t^i	$Princ. Comp1_t^i$	$Princ. Comp2_t^i$	$Princ. Comp3_t^i$
θ_{∞}^i	0.000 (0.000)	0.038 (0.209)	-0.011 (0.030)	-0.005 (0.033)	-0.005 (0.011)	-0.012 (0.023)	-0.012 (0.009)	0.001 (0.006)	-0.005 (0.017)	-0.008 (0.006)	0.002 (0.013)	-0.001 (0.006)	0.005 (0.025)	-0.001 (0.005)	0.003 (0.007)

Table 1.3: Parameter Estimates - Spain

Spain		Matrix Φ^{5p}														
VAR of Factors		IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$	IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$	
$X_t^{5p} = \Phi^{5p} X_{t-1}^{5p} + \Sigma^{5p} \epsilon_t^{5p}$		(0.057)	(0.088)	(0.068)	(0.038)	(0.095)	(0.032)	(0.033)	(0.087)	(0.073)	(0.056)	(0.069)	(0.029)	(0.027)	(0.029)	
IP^{5p}	-0.214***	0.372***	-0.133**	-0.039	0.029	0.105***	-0.13***	0.027	0.344***	0.078	-0.123**	0.279***	-0.113***	-0.064**	0.052*	
CP^{5p}	-0.066	0.161***	0.16***	0.077***	-0.064	0.058**	0.042**	-0.016	0.054	0.766***	0.114***	0.088**	-0.032	-0.032	-0.017	
Def^{5p}	(0.040)	(0.060)	(0.046)	(0.029)	(0.046)	(0.023)	(0.021)	(0.058)	(0.058)	(0.053)	(0.039)	(0.048)	(0.020)	(0.020)	(0.020)	
Deb^{5p}	0.212**	-0.138	0.086	0.086	0.474***	-0.027	0.076	0.148***	-0.258*	0.004	0.253***	0.221*	0.081	0.023	0.008	
$Princ. Comp1^{5p}$	(0.100)	(0.141)	(0.117)	(0.070)	(0.124)	(0.056)	(0.056)	(0.053)	(0.142)	(0.124)	(0.094)	(0.124)	(0.049)	(0.046)	(0.051)	
$Princ. Comp2^{5p}$	0.019	-0.015	-0.012	-0.009	0.010	-0.009	-0.001	0.009	0.001	-0.008	0.018	0.973***	0.010	0.005	-0.002	
$Princ. Comp3^{5p}$	(0.015)	(0.020)	(0.016)	(0.010)	(0.016)	(0.008)	(0.008)	(0.008)	(0.021)	(0.019)	(0.014)	(0.017)	(0.007)	(0.007)	(0.007)	
$Princ. Comp1^{5p}$	0.095	0.278	0.178	-0.001	0.214	-0.063	0.062	0.054	-0.198	-0.129	-0.035	0.221	0.418***	-0.208***	-0.115	
$Princ. Comp2^{5p}$	(0.153)	(0.228)	(0.181)	(0.103)	(0.186)	(0.085)	(0.081)	(0.081)	(0.223)	(0.193)	(0.144)	(0.192)	(0.074)	(0.073)	(0.071)	
$Princ. Comp3^{5p}$	-0.307***	-0.031	-0.120	-0.154	-0.110	0.129	0.065	0.14*	-0.140	0.226	0.117	0.101	-0.16**	0.473***	-0.088	
	(0.147)	(0.201)	(0.159)	(0.099)	(0.183)	(0.081)	(0.081)	(0.080)	(0.196)	(0.192)	(0.132)	(0.178)	(0.069)	(0.071)	(0.066)	
	0.124	-0.49**	0.114	0.101	-0.057	0.133	0.032	-0.082	0.694***	0.017	0.23*	-0.268	0.091	0.027	0.462***	
	(0.140)	(0.214)	(0.171)	(0.105)	(0.175)	(0.086)	(0.083)	(0.082)	(0.217)	(0.205)	(0.138)	(0.166)	(0.071)	(0.076)	(0.072)	
Short Rate Regression		Matrix Σ^{5p}														
$spread_t^{5p} = \delta^{5p} + \delta^{5p} X_{t-1}^{5p} + u_t^{5p}$		HY	IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$	IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$
$\delta_{\omega}^{5p} = \theta_{\omega}^{5p} + \theta_{\omega}^{5p} X_{t-1}^{5p}$		(0.001)	(0.023)	(0.027)	(0.032)	(0.047)	(0.029)	(0.030)	(0.030)	0.002	0.315***	-	-	-	-	-
δ_{ω}^{5p}		0.036***	-0.001	-0.001	-0.002	-0.071	-0.003	0.001	0.039	0.130	-	-	-	-	-	-
θ_{ω}^{5p}		(0.002)	(0.005)	(0.007)	(0.005)	(0.003)	(0.006)	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)	(0.005)	(0.003)	(0.002)	(0.002)
Recovery Intensity		Vector δ_{ω}^{5p}														
$\ln(\theta^{5p}) = \theta^{5p} + \theta^{5p} X_{t-1}^{5p}$		HY	IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$	IP^{5p}	CP^{5p}	Def^{5p}	Deb^{5p}	$Princ. Comp1^{5p}$	$Princ. Comp2^{5p}$	$Princ. Comp3^{5p}$
θ_{ω}^{5p}		0.000	-0.016	0.005	-0.009	0.006	0.004	-0.002	0.004	-0.003	0.017	-0.002	0.009	-0.007	-0.004	-0.006
		(0.004)	(0.549)	(0.027)	(0.032)	(0.034)	(0.027)	(0.027)	(0.017)	(0.020)	(0.029)	(0.017)	(0.021)	(0.013)	(0.023)	(0.014)

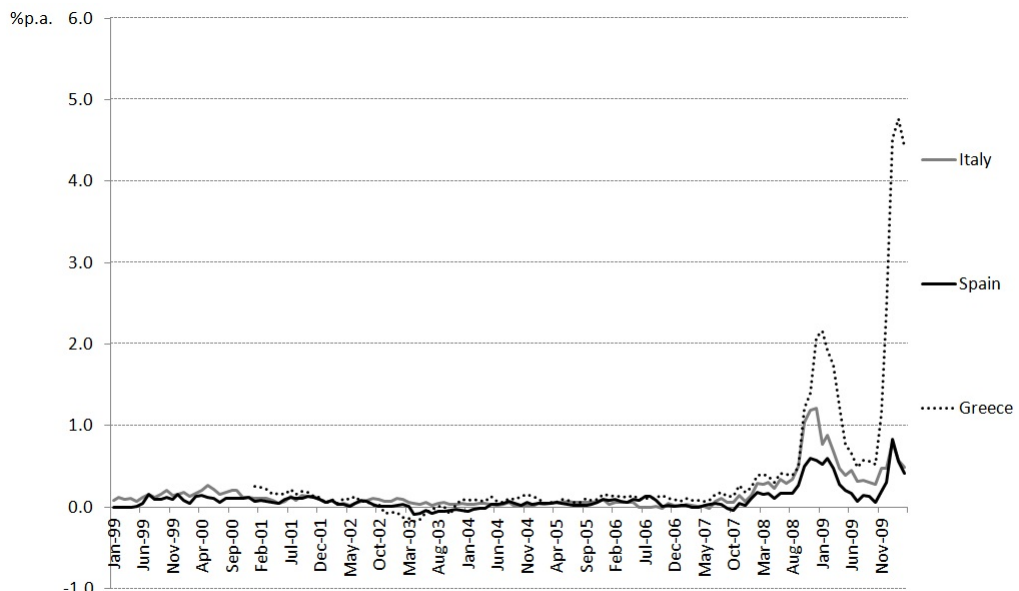
Table 1.4: Parameter Estimates - Greece

Greece		Matrix Θ^G													
VAR of Factors		HY_t	$IP_t^{G^a}$	$CPI_t^{G^a}$	$Def_t^{G^a}$	$Deb_t^{G^a}$	$Princ. Comp1_t^{G^a}$	$Princ. Comp2_t^{G^a}$	$Princ. Comp3_t^{G^a}$	$IP_t^{G^c}$	$Def_t^{G^c}$	$Deb_t^{G^c}$	$Princ. Comp1_t^{G^c}$	$Princ. Comp2_t^{G^c}$	$Princ. Comp3_t^{G^c}$
$X_t^{G^a} = \Theta^G X_{t-1}^{G^a} + \Sigma^G \epsilon_t^G$															
$IP_t^{G^a}$	-0.163	0.271***	-0.080	0.010	-0.119	0.199***	-0.245***	-0.024	-0.018	0.206**	-0.013	-0.35***	-0.061	-0.092*	0.104*
	(0.123)	(0.095)	(0.085)	(0.077)	(0.141)	(0.052)	(0.060)	(0.061)	(0.097)	(0.092)	(0.088)	(0.114)	(0.053)	(0.053)	(0.053)
$CPI_t^{G^a}$	-0.221***	0.15**	0.132**	0.031	-0.261***	0.029	0.041	0.033	-0.056	0.756***	0.038	0.224***	0.050	0.024	0.054
	(0.083)	(0.062)	(0.056)	(0.050)	(0.095)	(0.034)	(0.038)	(0.038)	(0.059)	(0.062)	(0.060)	(0.073)	(0.032)	(0.035)	(0.034)
$Def_t^{G^a}$	0.180	-0.224**	-0.043	-0.153*	0.593***	-0.087	-0.070	0.065	0.103	0.059	0.328***	-0.161	0.154***	0.035	-0.036
	(0.130)	(0.096)	(0.090)	(0.079)	(0.155)	(0.059)	(0.066)	(0.062)	(0.097)	(0.098)	(0.093)	(0.120)	(0.057)	(0.056)	(0.056)
$Deb_t^{G^a}$	0.027	-0.088***	0.024	0.014	0.064	-0.053***	0.008	0.034	0.035	-0.059**	0.013	0.916***	0.05***	-0.036**	-0.038**
	(0.037)	(0.030)	(0.027)	(0.023)	(0.046)	(0.016)	(0.020)	(0.019)	(0.030)	(0.029)	(0.028)	(0.035)	(0.017)	(0.016)	(0.016)
$Princ. Comp1_t^{G^a}$	-0.033	0.231**	-0.071	-0.035	-0.184	-0.032	0.062	-0.057	-0.122	0.059	0.236**	0.177	0.68***	-0.039	-0.006
	(0.138)	(0.107)	(0.096)	(0.080)	(0.164)	(0.063)	(0.070)	(0.066)	(0.107)	(0.105)	(0.100)	(0.139)	(0.063)	(0.058)	(0.059)
$Princ. Comp2_t^{G^a}$	-0.005	0.020	-0.053	-0.117	0.059	0.029	-0.098	0.143*	-0.019	0.121	0.003	0.006	0.010	0.586***	-0.163**
	(0.150)	(0.122)	(0.108)	(0.091)	(0.178)	(0.071)	(0.081)	(0.075)	(0.129)	(0.125)	(0.115)	(0.146)	(0.064)	(0.063)	(0.065)
$Princ. Comp3_t^{G^a}$	0.335**	0.444***	0.203*	0.23**	0.096	-0.223***	0.047	-0.225***	0.034	-0.308**	-0.057	0.262	0.087	-0.065	0.359***
	(0.165)	(0.132)	(0.120)	(0.102)	(0.197)	(0.083)	(0.080)	(0.082)	(0.138)	(0.122)	(0.129)	(0.161)	(0.069)	(0.072)	(0.075)
		Matrix Σ^G													
HY_t	$IP_t^{G^a}$	$CPI_t^{G^a}$	$Def_t^{G^a}$	$Deb_t^{G^a}$	$Princ. Comp1_t^{G^a}$	$Princ. Comp2_t^{G^a}$	$Princ. Comp3_t^{G^a}$	$IP_t^{G^c}$	$Def_t^{G^c}$	$Deb_t^{G^c}$	$Princ. Comp1_t^{G^c}$	$Princ. Comp2_t^{G^c}$	$Princ. Comp3_t^{G^c}$		
-0.002	-0.001	0.004	-0.002	-0.003	0.001	-0.002	0.003	0.468***	-	-	-	-	-		
(0.042)	(0.048)	(0.045)	(0.048)	(0.052)	(0.045)	(0.044)	(0.045)	(0.039)	-	-	-	-	-		
0.000	0.000	0.000	-0.001	-0.029	0.000	-0.002	-0.003	-0.031	0.266***	-	-	-	-		
(0.027)	(0.029)	(0.033)	(0.029)	(0.087)	(0.026)	(0.027)	(0.030)	(0.035)	(0.099)	-	-	-	-		
0.008	0.005	-0.003	0.000	-0.021	0.002	-0.003	-0.006	-0.204***	0.054	0.445***	-	-	-		
(0.042)	(0.046)	(0.048)	(0.049)	(0.050)	(0.041)	(0.049)	(0.046)	(0.056)	(0.057)	(0.049)	-	-	-		
-0.002	0.000	0.000	0.000	-0.001	0.000	0.002	0.000	-0.031	0.003	0.023	0.142***	-	-		
(0.013)	(0.014)	(0.014)	(0.015)	(0.015)	(0.013)	(0.015)	(0.015)	(0.019)	(0.016)	(0.014)	(0.014)	-	-		
	-	-	-	-	-	-	-	-	0.565***	-	-	-	-		
									(0.067)	-	-	-	-		
									0.021	0.615***	-	-	-		
									(0.080)	(0.088)	-	-	-		
									-0.091	0.140	0.62***	-	-		
									(0.089)	(0.154)	(0.079)	-	-		
Short Rate Regression		Vector $\delta_t^{G^c}$													
$spread_t^{G^a} = \delta_t^{G^a} + \delta_t^{G^c} X_{t-1}^{G^a} + u_t^{G^a}$	HY_t	$IP_t^{G^a}$	$CPI_t^{G^a}$	$Def_t^{G^a}$	$Deb_t^{G^a}$	$Princ. Comp1_t^{G^a}$	$Princ. Comp2_t^{G^a}$	$Princ. Comp3_t^{G^a}$	$IP_t^{G^c}$	$Def_t^{G^c}$	$Deb_t^{G^c}$	$Princ. Comp1_t^{G^c}$	$Princ. Comp2_t^{G^c}$	$Princ. Comp3_t^{G^c}$	
0.301***	-0.009	0.384***	0.253***	-0.054***	0.031***	0.298***	-0.205***	-0.057***	0.032***	0.122***	0.166***	0.315***	0.137***	-0.029***	
(0.002)	(0.005)	(0.004)	(0.004)	(0.003)	(0.006)	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.002)	(0.002)	
Recovery Intensity		Vector $\theta_t^{G^c}$													
$\ln(\theta_t^{G^c}) = \theta_t^{G^c} + \theta_t^{G^c} X_{t-1}^{G^a}$	HY_t	$IP_t^{G^a}$	$CPI_t^{G^a}$	$Def_t^{G^a}$	$Deb_t^{G^a}$	$Princ. Comp1_t^{G^a}$	$Princ. Comp2_t^{G^a}$	$Princ. Comp3_t^{G^a}$	$IP_t^{G^c}$	$Def_t^{G^c}$	$Deb_t^{G^c}$	$Princ. Comp1_t^{G^c}$	$Princ. Comp2_t^{G^c}$	$Princ. Comp3_t^{G^c}$	
0.000	-0.009	0.006	-0.003	0.000	-0.002	0.012	-0.002	-0.001	-0.005	0.008	0.002	-0.008	0.000	-0.012	
(0.001)	(1.413)	(0.104)	(0.066)	(0.049)	(0.108)	(0.084)	(0.029)	(0.071)	(0.044)	(0.129)	(0.035)	(0.058)	(0.029)	(0.028)	

Table 1.5: Autocorrelation of yields, spreads and the three first principal components (PC1, PC2 and PC3) of yields (or spreads, for the IGS countries) orthogonalized with respect to the macroeconomics and risk factors

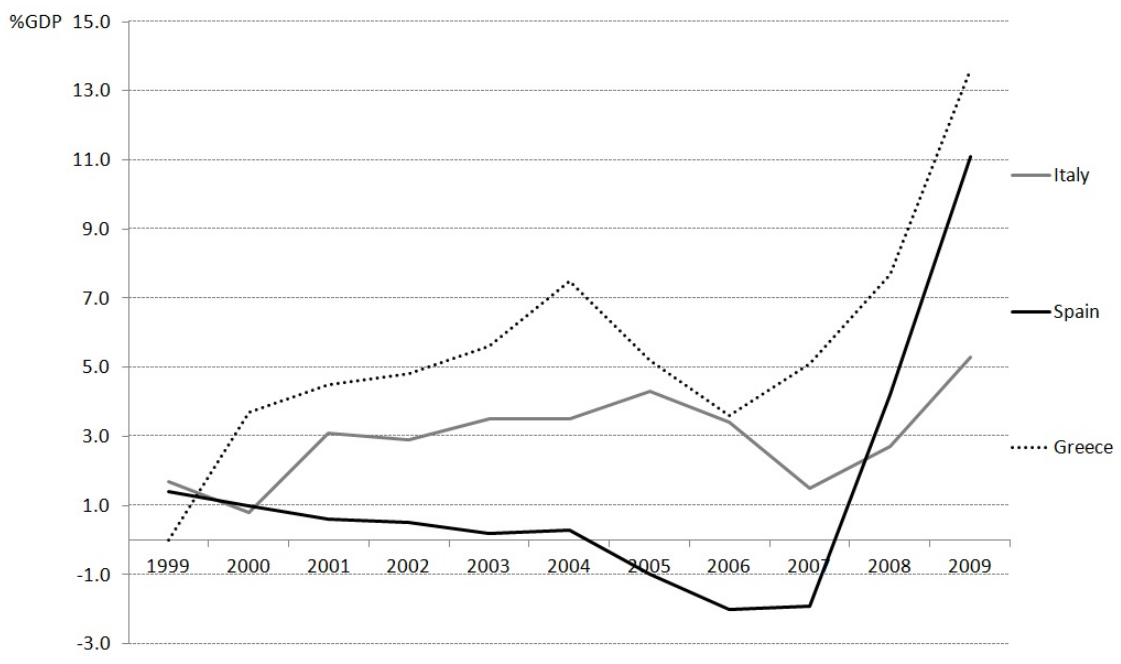
Lags	Yields(Germany) and Spreads(Spain, Italy and Greece) Maturities									PC1	PC2	PC3
	3	6	12	24	36	48	60	84	120			
<i>Germany</i>												
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.97	0.97	0.97	0.96	0.95	0.95	0.95	0.95	0.95	0.74	0.61	0.73
2	0.93	0.92	0.92	0.90	0.89	0.88	0.88	0.89	0.89	0.49	0.46	0.58
3	0.87	0.87	0.86	0.83	0.83	0.82	0.82	0.84	0.84	0.28	0.39	0.49
<i>Spain</i>												
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.77	0.78	0.87	0.91	0.94	0.95	0.95	0.94	0.96	0.54	0.45	0.07
2	0.58	0.62	0.73	0.82	0.87	0.89	0.90	0.89	0.92	0.29	0.09	0.05
3	0.47	0.56	0.61	0.70	0.78	0.81	0.84	0.81	0.86	0.21	-0.09	0.07
<i>Italy</i>												
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.66	0.61	0.77	0.86	0.91	0.92	0.94	0.93	0.95	0.40	0.47	0.50
2	0.57	0.53	0.57	0.69	0.78	0.79	0.84	0.84	0.90	0.11	0.19	0.21
3	0.52	0.47	0.41	0.51	0.62	0.65	0.72	0.74	0.83	-0.05	0.05	0.12
<i>Greece</i>												
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	0.81	0.82	0.83	0.84	0.87	0.88	0.89	0.89	0.90	0.65	0.57	0.35
2	0.55	0.55	0.57	0.61	0.68	0.70	0.73	0.73	0.77	0.27	0.27	-0.02
3	0.30	0.29	0.33	0.37	0.46	0.49	0.54	0.54	0.62	0.01	-0.09	-0.11

Figure 1.1: 1 Year Sovereign Bond Yield Spread (%p.a.) - Selected Countries



Source: Eurostat

Figure 1.2: Central Government Budget Deficit



Source: Eurostat

Figure 1.3: Deficit vs. 1 Year Bond Yield Spreads - 4 quarters Moving Average

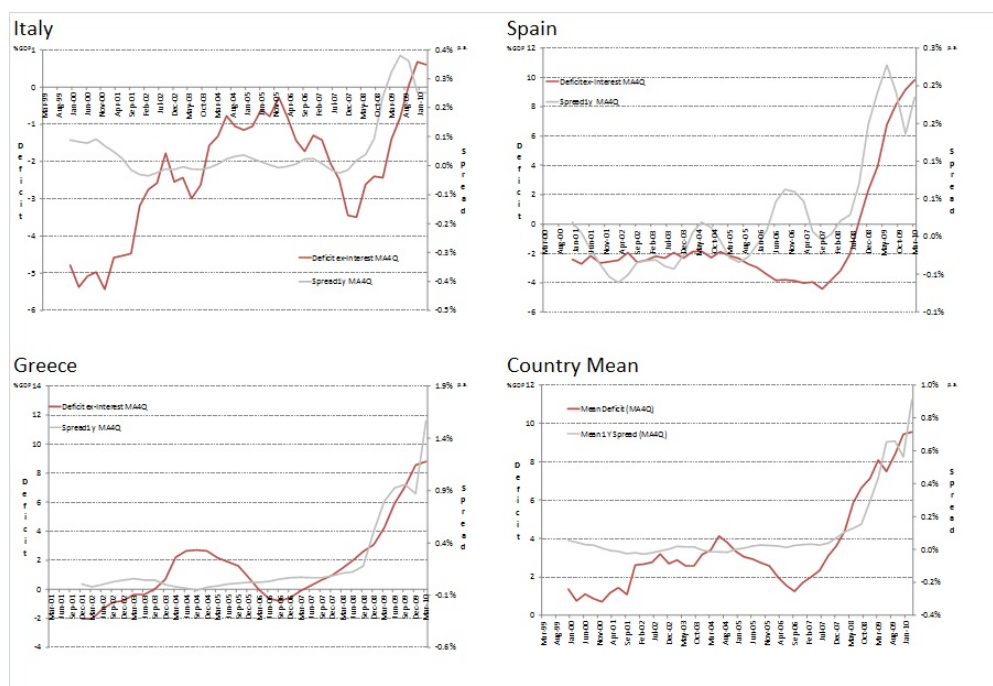


Figure 1.4: Moody's High Yield (x-axis) and 10-year bond yields, scatter plots

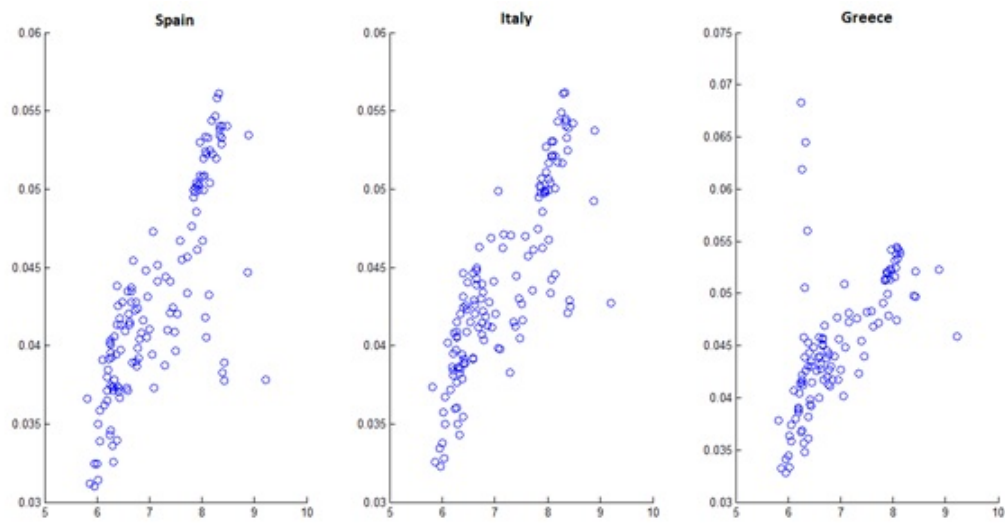
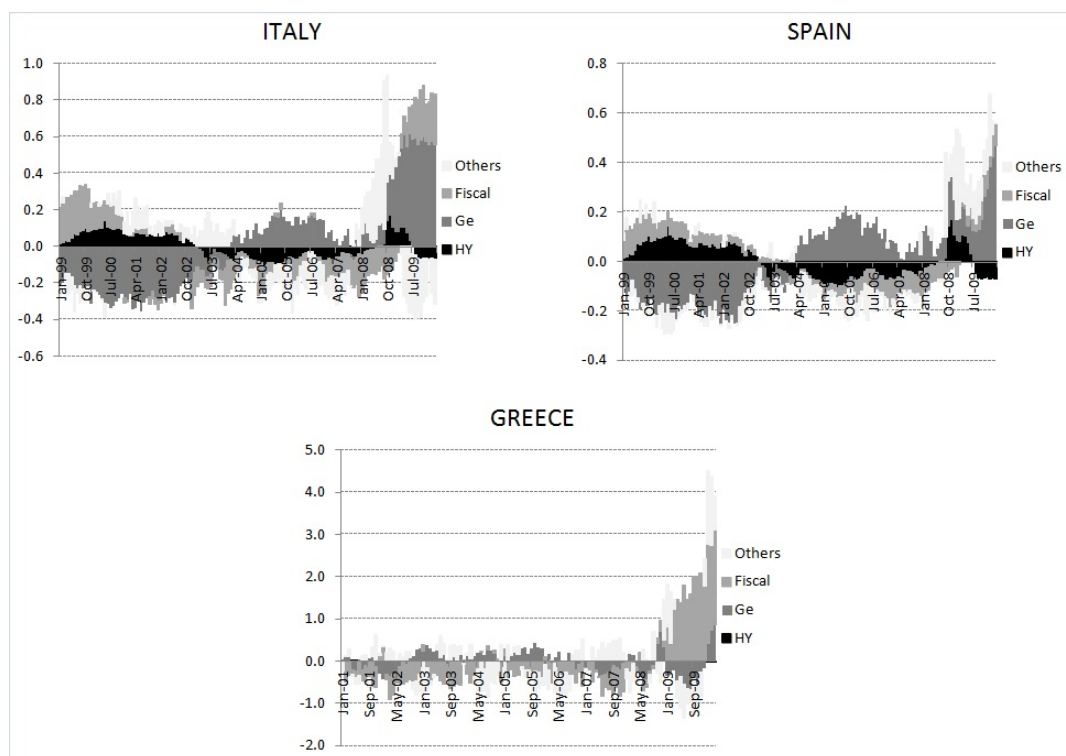
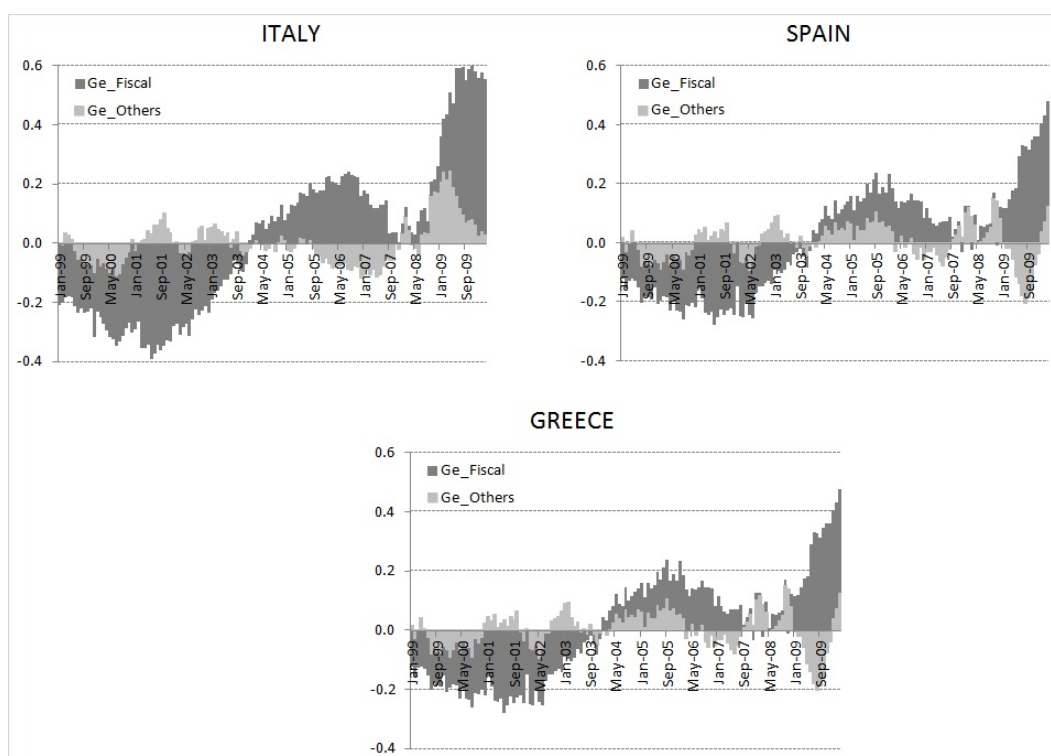


Figure 1.5: 1 year bond yield spread estimated composition



Composition based on the E^N factor weights in the 1 year bond yield spread (i.e., for $N=12$) multiplied by the current values of the corresponding factor along time. The factors *Others* are IP, Inf and the principal components of the spreads (summed up in PC). Factors *Ge are* Germany related variables. Factors *Fiscal* are Def and Deb . Y-axis scales are different among countries for a better visualization.

Figure 1.6: 1 year bond yield spread estimated composition - Only Germany related factors



Composition based on Germany related factors. Composition based on the E^N factor weights in the 1 year bond yield spread (i.e., for $N=12$) multiplied by the current values of the corresponding factor along time. Factors Other are Ge_{IP} , Ge_{Inf} , and the principal components of the Germany's yields (Ge_{1PC} to Ge_{3PC}). Factors Ge_{Fiscal} are Ge_{Def} , Ge_{Deb} .

Figure 1.7: E^N factor loadings of HY , Ge_Deb , IP , Inf , Def and Deb for the three countries as a function of maturity N .

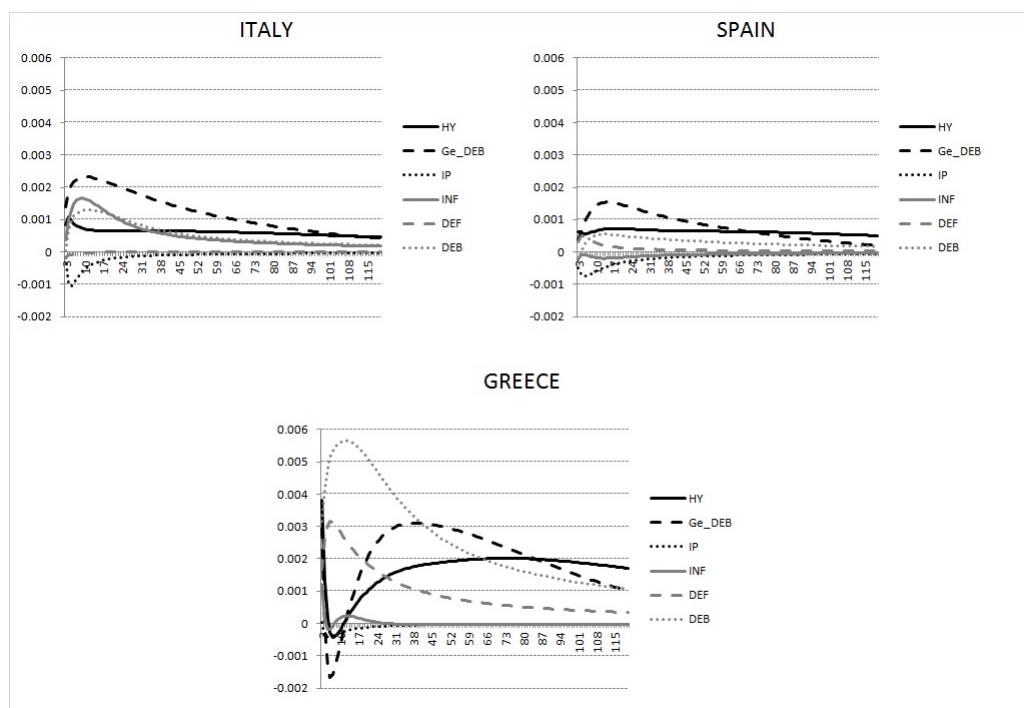
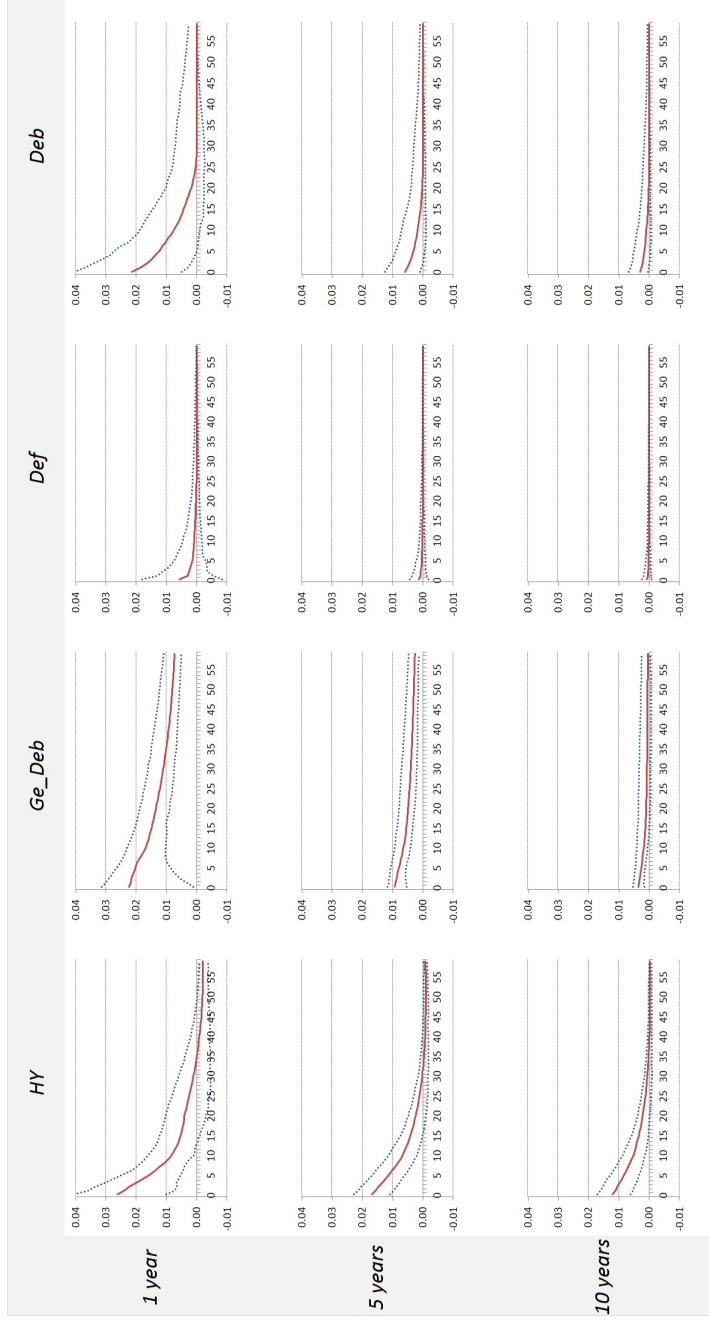
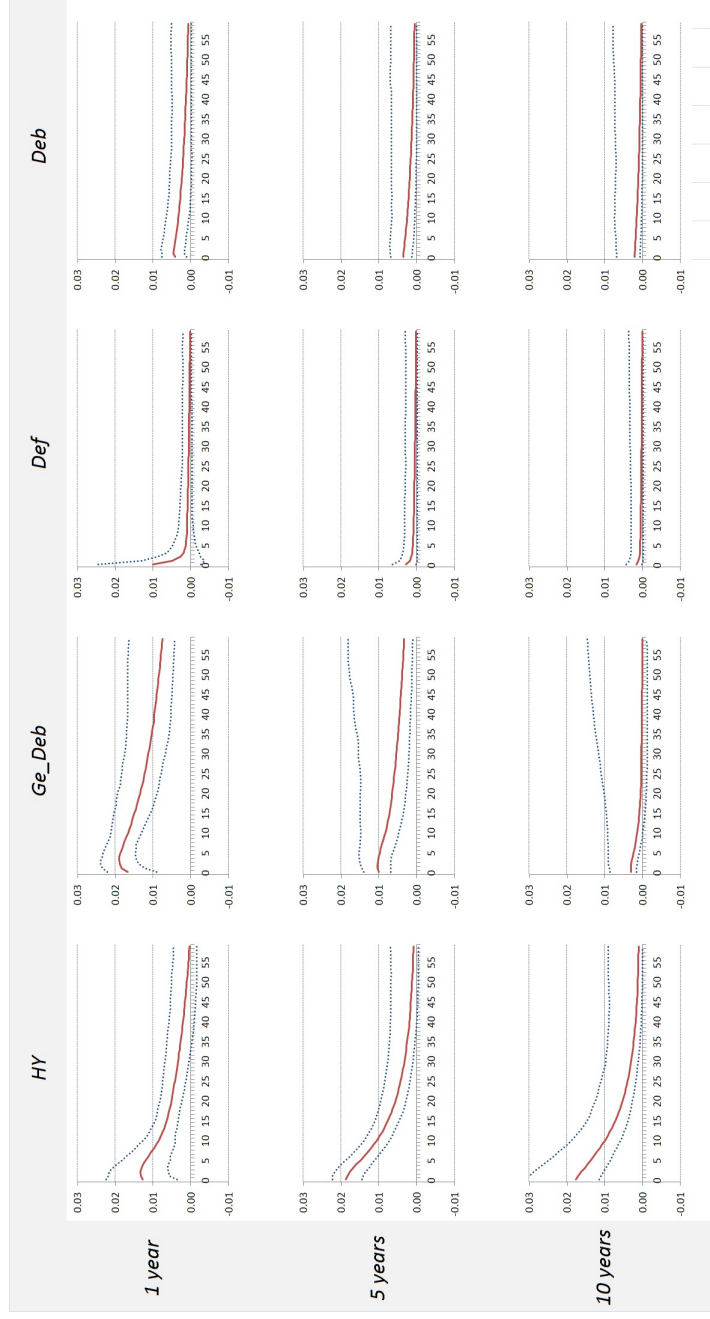


Figure 1.8: Impulse response functions for Italy.



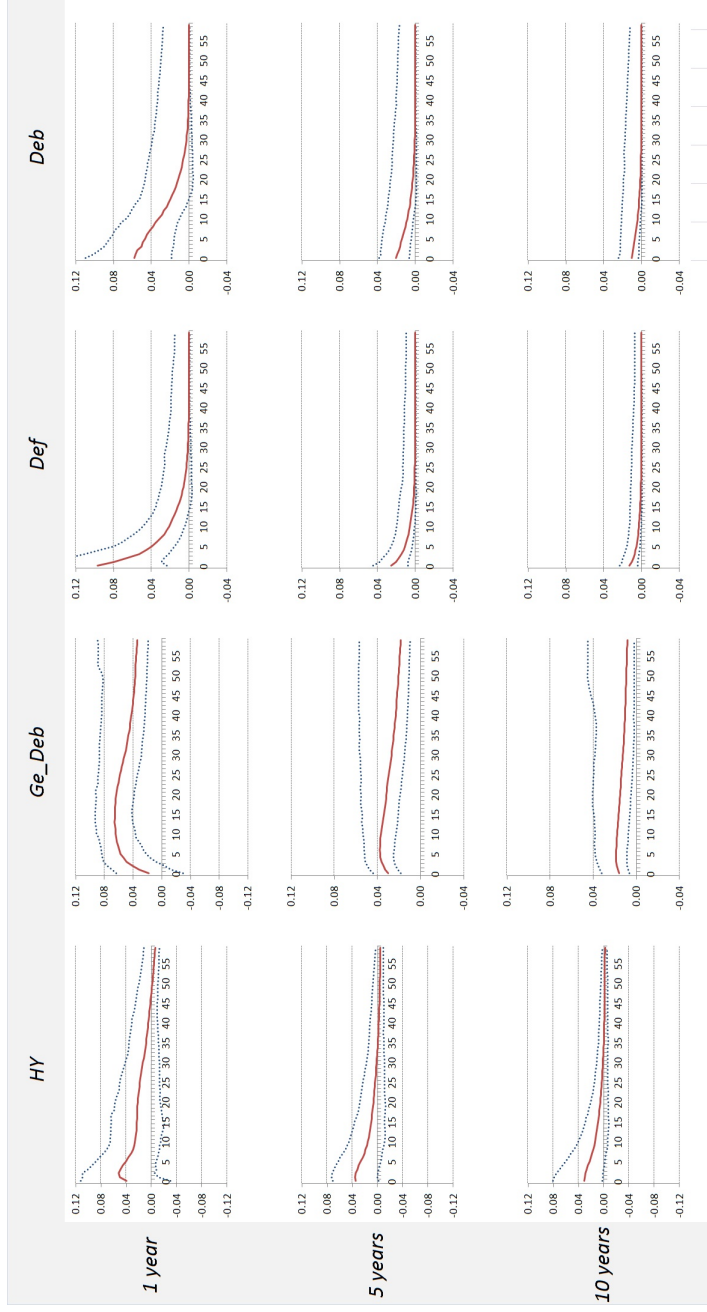
All Impulse Responses (IR's) are from a one standard deviation shocks in HY, Ge_Deb, Def and Deb. IR's for 12 month (top row), 60 month (middle row) and 120 month (bottom row) yield spreads. Confidence intervals of 95%. X-axis scales are different among countries for better visualization. Responses are multiplied by 100.

Figure 1.9: Impulse response functions for Spain.



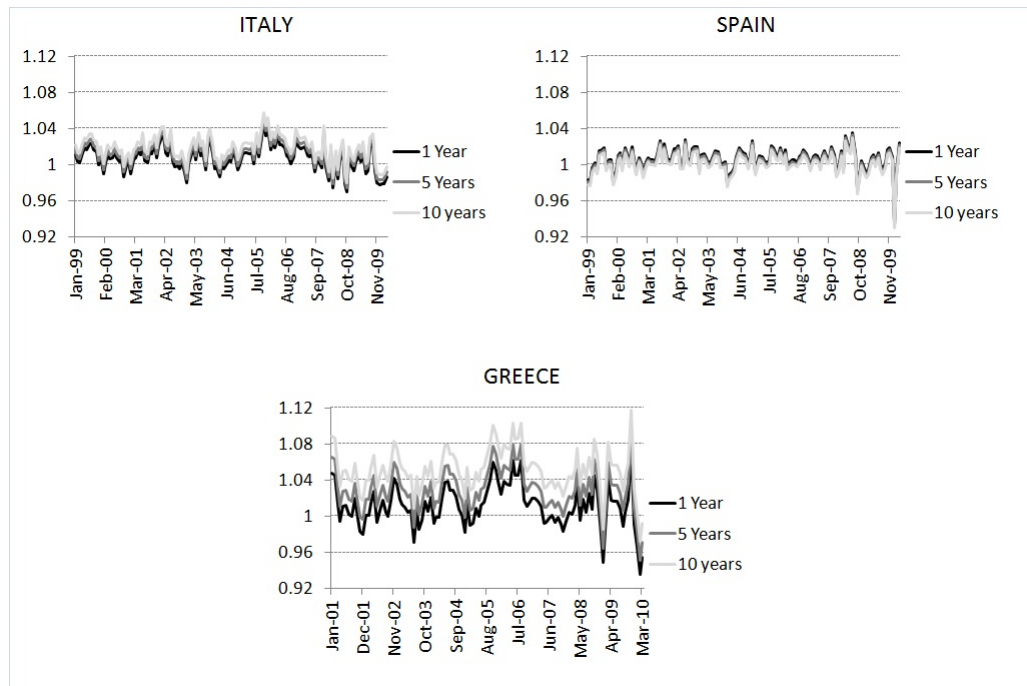
All Impulse Responses (IR's) are from a one standard deviation shocks in HY, Ge_Deb, Def and Deb. IR's for 12 month (top row), 60 month (middle row) and 120 month (bottom row) yield spreads. Confidence intervals of 95%. X-axis scales are different among countries for better visualization. Responses are multiplied by 100.

Figure 1.10: Impulse response functions for Greece.



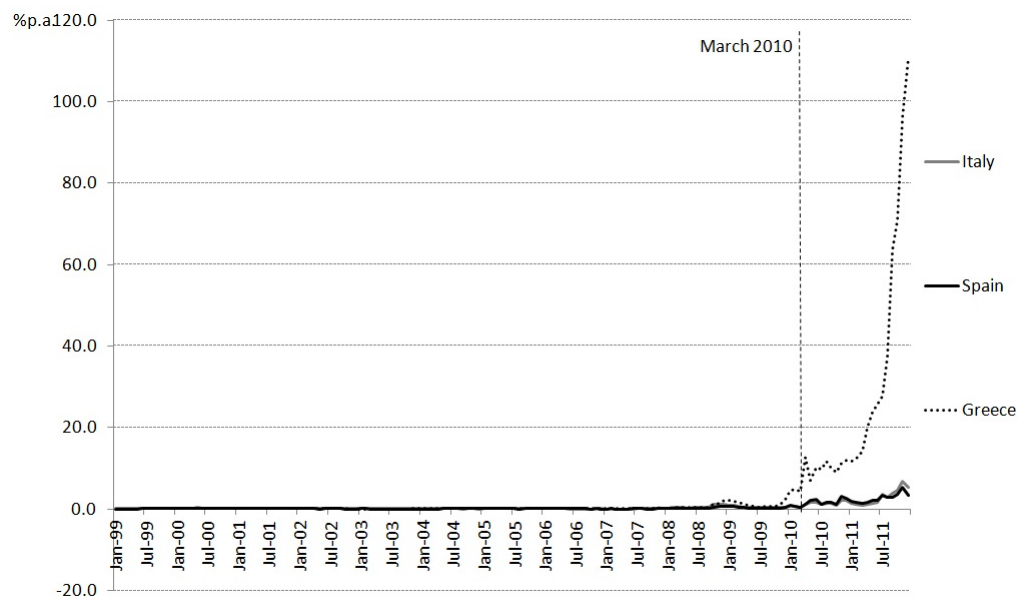
All Impulse Responses (IR's) are from a one standard deviation shocks in HY, Ge_Deb, Def and Deb. IR's for 12 month (top row), 60 month (middle row) and 120 month (bottom row) yield spreads. Confidence intervals of 95%. X-axis scales are different among countries for better visualization. Responses are multiplied by 100.

Figure 1.11: Path of Θ_t^N for maturities 1, 5 and 10 years for Italy, Spain and Greece



$\ln\Theta_t^N$ follows the process $\ln\Theta_t^N = \theta_t^{iN} = \theta_{00}^i + N\theta_{01}^i + \theta_1^{i'} \mathbf{X}_t^i$, where N is the maturity and \mathbf{X}_t^i is Germany's and country i 's factors.

Figure 1.12: 1 Year Sovereign Bond Yield Spread (%p.a.) - Italy, Spain and Greece. Sample: January 1999 to December 2011



Source: Eurostat

Figure 1.13: 1 Year Sovereign Bond Yield Spread (%p.a.) - Italy, Spain and Greece.
Sample: January 1999 to December 2011

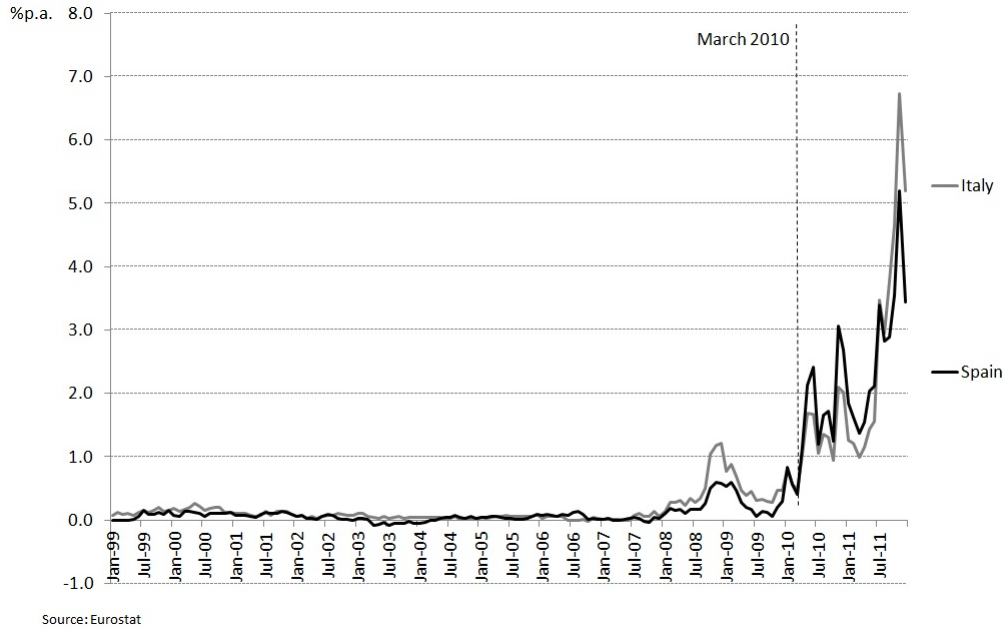


Figure 1.14: E^N factor loadings of HY , Ge_{Deb} , IP , $Infl$, Def , Deb for Italy and Spain as a function of maturity N - Sample: January 1999 to December 2011

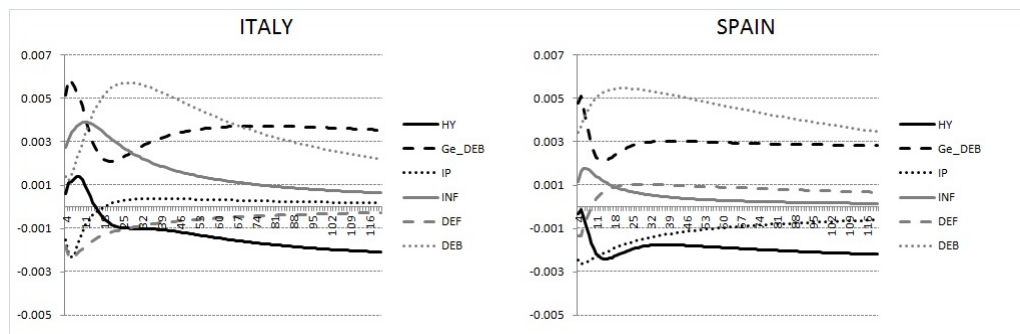


Figure 1.15: Deficit (ex-interest rate payments) as a percentage of GDP

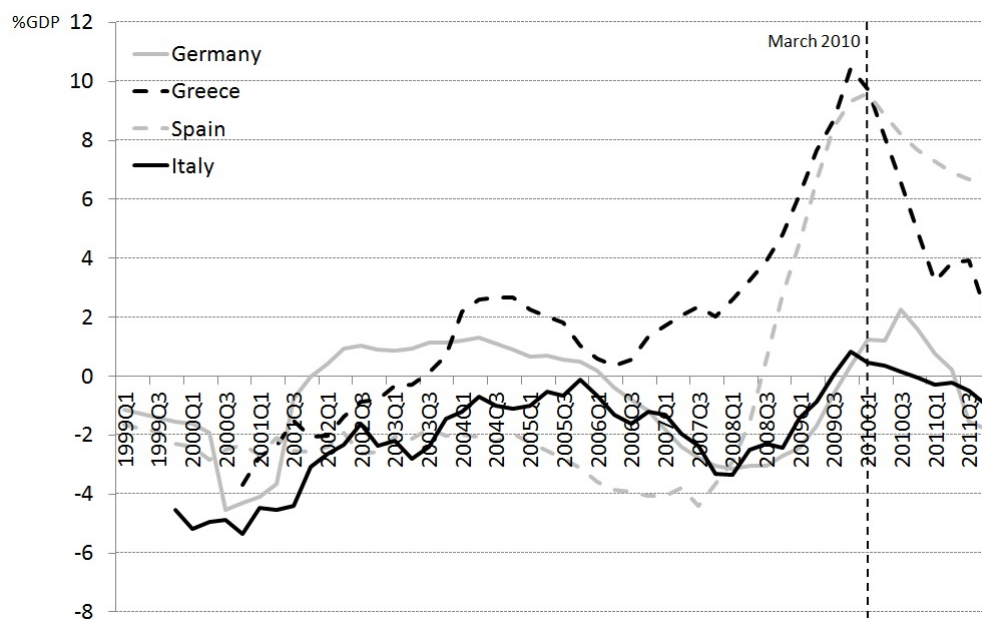
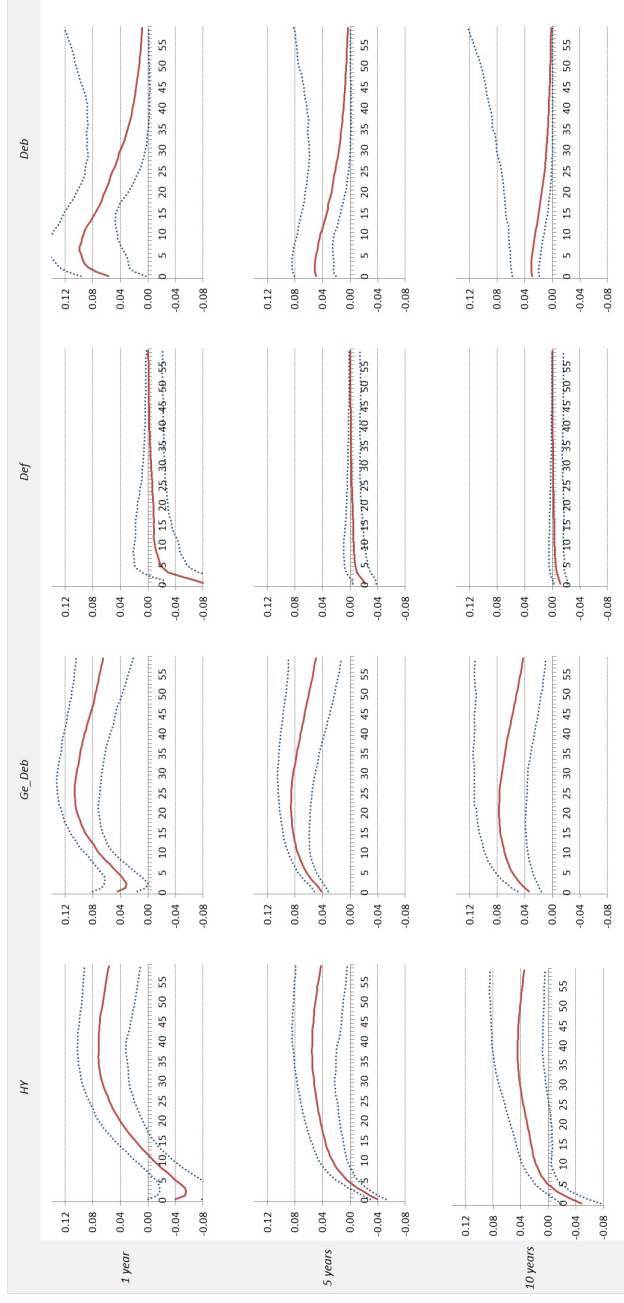
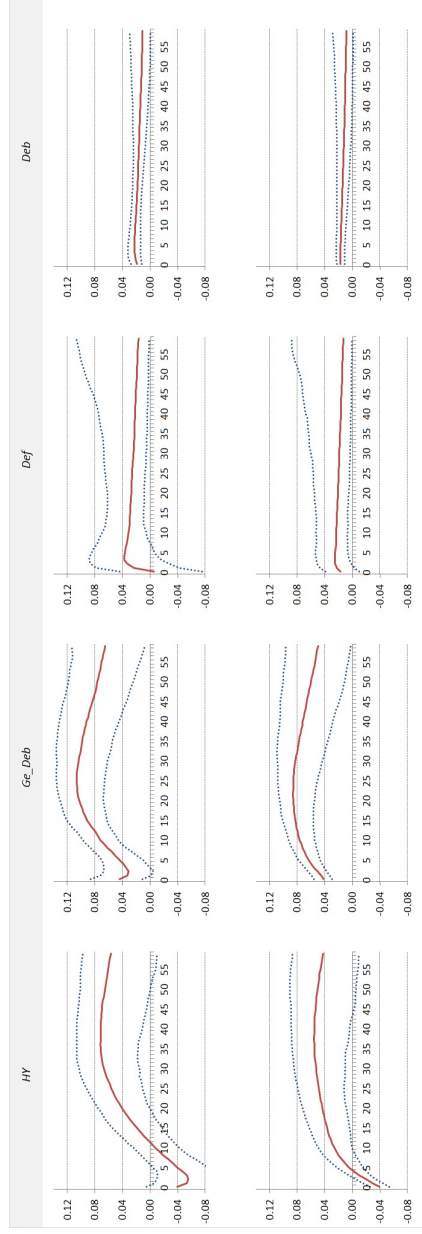


Figure 1.16: Impulse response functions for Italy - Sample: January 1999 to December 2011



All Impulse Responses (IR's) are from a one standard deviation shocks in HY, Ge_De, Def and Deb. IR's for 12 month (top row), 60 month (middle row) and 120 month (bottom row) yield spreads. Confidence intervals of 95%. Responses are multiplied by 100.

Figure 1.17: Impulse response functions for Spain - Sample: January 1999 to December 2011

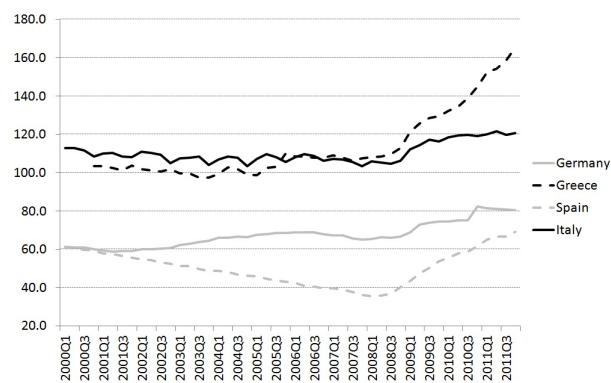


All Impulse Responses (IR's) are from a one standard deviation shocks in HY, Ge_De, Def and Deb. IR's for 12 month (top row), 60 month (middle row) and 120 month (bottom row) yield spreads. Confidence intervals of 95%. Responses are multiplied by 100.

Table 1.6: Debt to GDP Ratio from 2000Q1 to 2010Q3 - Granger Causality Tests

Null Hypothesis	Lag=1		Lag=2		Lag=3		Lag=4	
	Obs	Probability	Obs	Probability	Obs	Probability	Obs	Probability
GREECE does not Granger Cause GERMANY	37	0.24	36	0.01***	35	0.02**	34	0.08*
GERMANY does not Granger Cause GREECE		0.22		0.16		0.11		0.29
ITALY does not Granger Cause GERMANY	40	0.72	39	0.43	38	0.25	37	0.70
GERMANY does not Granger Cause ITALY		0.15		0.09*		0.06*		0.01***
SPAIN does not Granger Cause GERMANY	40	0.82	39	0.03**	38	0.03**	37	0.11
GERMANY does not Granger Cause SPAIN		0.00***		0.54		0.13		0.07*
ITALY does not Granger Cause GREECE	37	0.35	36	0.46	35	0.72	34	0.19
GREECE does not Granger Cause ITALY		0.00***		0.00***		0.01***		0.0***
SPAIN does not Granger Cause GREECE	37	0.27	36	0.00***	35	0.01***	34	0.04**
GREECE does not Granger Cause SPAIN		0.00***		0.75		0.90		0.95
SPAIN does not Granger Cause ITALY	40	0.88	39	0.00***	38	0.00***	37	0.00***
ITALY does not Granger Cause SPAIN		0.00***		0.56		0.53		0.09

Figure 1.18: Debt to GDP ratio from 2000Q1 to 2010Q3



2 Interest and Exchange Rates Integrated Model

2.1 Introduction

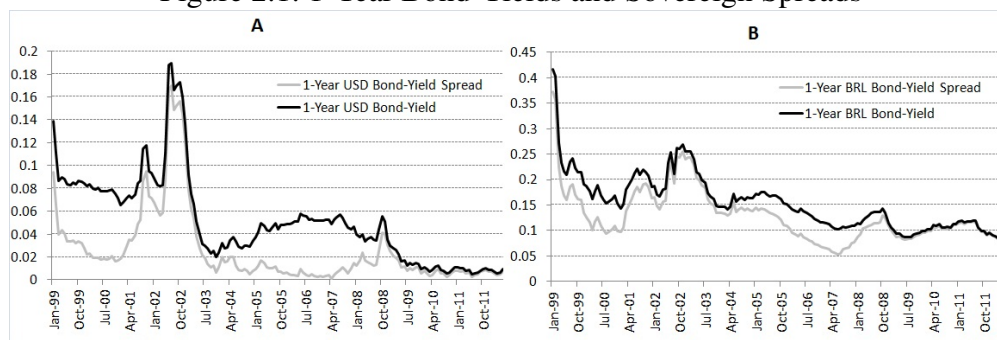
In light of a model that ensures the attendance of no-arbitrage restrictions, we investigate the determinants of the term structure of yields and the exchange rate expectations in Brazil. Our goal is to explain the general trend of reduction in Brazilian sovereign spreads observed in the last seven years in both USD and Brazilian real (BRL) bond market. The path of 1-year bond yields and spreads can be seen in Figures 1A and 1B. We can see that yields and spreads reached their lowest levels in recent years. Our strategy is the following: (i) through the interaction between the Brazilian and the American USD bond-yield curves, we capture the default risk on the Brazilian USD-denominated bond; (ii) through the interaction between the Brazilian BRL and USD bond-yield curves, we capture the exchange rate expectations.

The exchange rate is defined as the amount BRL needed to purchase 1 USD, while interest rates for various periods (or maturities) may involve risk or not. If there is no risk, these rates are drawn from the market prices of zero coupon bonds issued by U.S. government, while in case with default risk, these rates are drawn from the market prices of zero coupon bonds issued by the Brazilian government and denominated in USD and in BRL. It is assumed that the securities issued by Brazilian government are risky, i.e., there is a likelihood that the issuer does not fulfill its obligations in contract (default).

Our model dialogues with three strands of literature on Macroeconomics and Finance. The first consists of models that examine theoretical and empirical curves of yields on bonds issued by firms with different risk levels who work in the same market (mostly the U.S.). The probability of an issuer of a security to face difficulties that prevent it from honoring all or part of its commitments makes the equilibrium price of a given security be adjusted so that investors receive an extra pay to compensate them for the default risk. There are many papers that study the yield curve of a defaultable bond, and among them we highlight Duffee (1996,1998), Duffie and Singleton (1999, 2003), and Amato and Luisi (2006).

In the second strand, the issuer is an emerging country and bonds are traded in the international market. The case in which the emerging country issues bonds

Figure 2.1: 1-Year Bond-Yields and Sovereign Spreads



Note: A yield of $x\%$ per year enters as $z = \log(1+x/100)$. If the yield on the Brazilian bond is $xbr\%$ per year and the yield on the U.S. bond is $xus\%$ per year, the spread enters as $s=zbr-zus$, where $zbr = \log(1+xbr/100)$ and $zus = \log(1+xus/100)$.

in an internationally accepted currency (for example, USD) in order to mitigate the exchange rate risk assumed by the investor was discussed in Bonomo and Lowenkron (2008) and in Wu and Zhang (2008). Other important references are Andritzky (2006), whose book covers in detail the main aspects related to sovereign debt, and Meres and Almeida (2008), that model jointly the yield curve of Brazilian and U.S. bonds taking into account the default risk inherent in the first one. The background of these papers is the literature on term structure models [Duffee (2002), Dai and Singleton (2000), among others]. Ang and Piazzesi (2003) introduced macroeconomic variables into Gaussian term structure models, leading to an explosion of macro-finance research. As noted by Duffee (2011), this literature is not designed to produce more accurate term structure models, but rather to explicitly link the term structure to its fundamental determinants.

More recently, Joslin, Singleton and co-authors (see Joslin, Priebsch and Singleton [2010] and Joslin, Le and Singleton [2012]) stressed that pricing strategies like Ang and Piazzesi (2003) imply that macroeconomic variables are spanned by model-implied bond yields, and this fact would not be supported by data. In our paper, the variables other than macroeconomic ones are constructed in a way that the macroeconomic variables are naturally spanned by them. It is also due to Joslin, Singleton and co-authors some evidence that a canonical macro-finance term structure model does not offer any new insights into economic questions regarding the historical distribution of the macro risk factors and yields, over and above what is learned from a factor-VAR (see Joslin, Le and Singleton [2012]). In the case of a model for the term structure of yields on defaultable bonds, the imposition of no-arbitrage restrictions allows us to recover the parameters that drive the total discount rate of the security due to default risk and exchange rate expectations.

Our model also dialogues with a third strand of the literature on Macroeconomics and Finance, which explores the link between interest rates on the various

maturities for developed countries and the variation of the nominal bilateral exchange rate. This literature uses an integrated affine model to describe the behavior of the two yield curves and turn the relationship between interest and exchange rates endogenous, reflecting the fundamentals of the model. The theoretical and empirical implications of this model (which belongs to the so called "Internationally Affine Term Structure Models") are studied in several articles, and among them we highlight Backus, Foresi and Telmer (2001), Anderson, Hammond and Ramezani (2010) and de los Rios (2009). It should be noted that the possibility of default is not taken into account by these articles, because the pairs of countries under study (USA and United Kingdom, USA and Canada, etc.) are allegedly immune to this risk.

The aim of our paper is more about understanding the fundamentals behind the reduction in Brazilian sovereign spreads observed in the last years in both USD and BRL bond market rather than promoting a particular modelling innovation. Our model has a core critical simplifying assumption: we assume that the default probability is the same for bonds denominated in USD and in BRL issued by the Brazilian government, and the difference in yields on these two kinds of bonds is due to exchange rate risk. Our exercises indicate that factors linked to the Brazilian economy had a positive contribution for the spreads from 1999 to 2006 and, in most of this period, factors linked to the U.S. economy have a negative contribution. From the end of 2007 on, this scenario clearly inverts. Additionally, we present evidence that the improving in our external solvency indicator (which takes into account external debt, current account flows, terms of trade, international price of commodities, among others), along with the improvement in the inflation scenario, is the main responsible for the reduction of spreads from the middle of 2007 on. We also found evidence that the expected discount on the price of the defaultable BRL bond due specifically to exchange rate risk does not vary with economic fundamentals or even with yields-related variables.

The reminder of this paper is organized as follows. In the second section, we discuss the underlying model and main assumptions. In the third section, we present data, and the results are discussed in the forth section. Section five concludes.

2.2 Underlying model and main assumptions

Our model is estimated is three steps. In a first step, a standard affine term structure model *a la* Ang and Piazzesi (2003) is estimated for the U.S. economy. In a second step, we estimate an affine term structure model for the USD-denominated bonds issued by the Brazilian government, accounting for default risk, using a framework similar to the one in Bonomo and Lowenkron (2008). In a third step, we estimate an affine term structure model for bonds denominated in BRL issued by the Brazilian government taking into account, besides the default risk estimated

in the second step, the exchange rate risk.

It is always important to stress that our model has the critical assumption that the default probability and liquidity premium is the same for bonds denominated in USD and in BRL issued by the Brazilian government, and the difference between yields on these two kinds of bonds is due to exchange rate risk. This assumption is made because it is virtually impossible to distinguish default risk from exchange rate movements risk when we observe the term structure of BRL-denominated bond-yields. Default risk and exchange rate movements risk are closely related and it is hard to sell that there are some factors that has an important role in the determination of only one of them. On the other hand, it is a bit easier to believe that we are able to distinguish default risk from exchange rate movements risk when we observe the term structure of USD-denominated bond-yields. This last curve would be affected by exchange rate risk if investors believed that the Brazilian national currency could reach a such low value that would make difficult for the Brazilian government to fulfill it obligations in a foreign currency. Observing the path of both nominal and real exchange rate BRL vs USD (see figure 2.3), it seems that this risk could have been of substantial size only around the end of 2002, when a confidence crisis emerged in country when the former president Lula was about to be elected. We also make the assumption that investors are not discriminated according to their nationality, so we can think of the buyer of the bond as a representative international investor.

The term structure model we estimate for the U.S. term structure of yields is standard. The short rate r_t follows an affine function of variables related to the U.S. economy grouped in the vector \mathbf{X}_t^{US} of dimension K^{US} (Duffie and Kan [1996]):

$$r_t = \delta_0 + \delta_1' \mathbf{X}_t^{US}. \quad (2-1)$$

The composition and dynamics of \mathbf{X}_t^{US} will be discussed bellow. No-arbitrage guarantees that the price at t of a $(N + 1)$ -period bond, V_t^{N+1} , is given by:

$$V_t^{N+1} = \mathbb{E}[M_{t+1} V_{t+1}^N], \quad (2-2)$$

where M_{t+1} is the the stochastic discount rate (or pricing kernel) at $t + 1$. M_{t+1} can be written as:

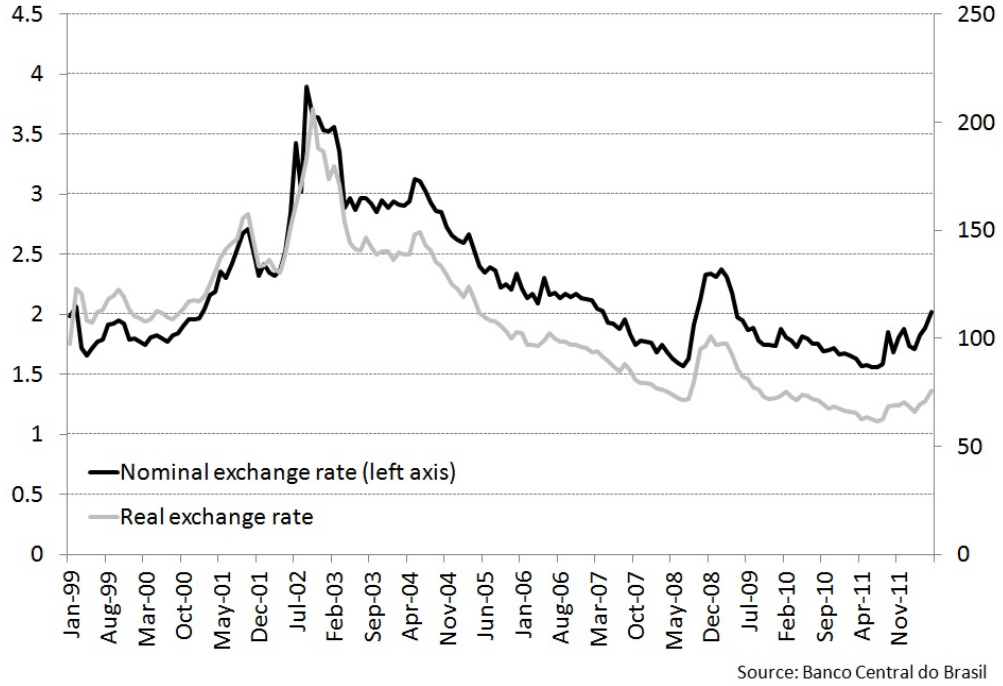
$$M_{t+1} = \exp\{-1/2\lambda_t' \lambda_t - \delta_0 - \delta_1' \mathbf{X}_t^{US} - \lambda_t' \varepsilon_{t+1}\}, \quad (2-3)$$

where λ_t is the market price of risk (time variant) associated with the sources of uncertainty ε_{t+1} . λ_t is parameterized as an affine function of the variables:

$$\lambda_t = \lambda_0 + \lambda_1' \mathbf{X}_t^{US}, \quad (2-4)$$

for a $K^{US} \times 1$ vector λ_0 and a $K^{US} \times K^{US}$ matrix λ_1 . The price at t of a $(N + 1)$ -period

Figure 2.2: Nominal and Real BRL vs USD exchange rate.



Source: Banco Central do Brasil

Note: The nominal exchange rate is the amount of BRS needed to buy one USD in the end of each month. The real exchange rate is an index (Jun/1994=100) that adjust the amount of BRS needed to buy one USD by the Brazilian consumer price index IPCA.

bond issued by the U.S. government can be written as:

$$V_t^{N+1,US} = \exp(A_{N+1} + B'_{N+1} X_t^{US}), \quad (2-5)$$

where A_{N+1} and B_{N+1} satisfy well-known Riccati difference equations (see Appendix F for a summary).

For the pricing of the USD-denominated bond issued by the Brazilian government (that henceforward will receive the superscript $BRUS$), we follow Bonomo and Lowenkron (2008) and write:

$$V_t^{N+1,BRUS} = \mathbb{E}[M_{t+1} \Theta_{t+1} V_{t+1}^{N,BRUS}], \quad (2-6)$$

where Θ_{t+1} , which we will call the *recovery intensity*, is a measure of default and liquidity risk. In our model, $\Theta_{t+1} = 1 - h_t L_{t+1} - l_{t+1}$, where h_t is probability at t of default between t and $t + 1$, L_{t+1} denotes the proportional loss in case of default at $t + 1$, and l_{t+1} is the liquidity risk at $t + 1$. We assume that $\ln \Theta_{t+1}$ is linear on factors X_t^{BR} which represents a K -dimensional vector of variables of interest for the Brazilian economy. Therefore, $\ln \Theta_t := \theta_t = \theta_0 + \theta'_1 \mathbf{X}_t^{BR}$.

As the Brazilian economy is peripheral, we assume that the K^{US} -dimensional vector X_t^{US} is a subset of the K -dimensional vector X_t^{BR} , i.e.,

$$\mathbf{X}^{BR} = [\mathbf{X}^{US} \quad \mathbf{X}^{BR*}],$$

where \mathbf{X}^{BR*} represents variables that are specific about the Brazilian economy, such as activity and inflation in Brazil. Naturally, $K^{US} \leq K$. We assume that the normalized (mean zero, variance one) \mathbf{X}_t^{BR} follows a Gaussian VAR (1). The VAR order was chosen by the Schwarz information criterion in a range between 0 and 6.

$$\mathbf{X}_t^{BR} = \Phi \mathbf{X}_{t-1}^{BR} + \Sigma^{BR} \varepsilon_t^{BR}. \quad (2-7)$$

Also because of the peripheral condition of the Brazilian economy, we assume that \mathbf{X}_t^{US} follows an independent dynamics, which yields

$$\mathbf{X}_t^{BR} = \begin{bmatrix} \mathbf{X}_t^{US} \\ \mathbf{X}_t^{BR*} \end{bmatrix} = \begin{bmatrix} \Phi^{US} & \mathbf{0} \\ \Phi^{BR*,US} & \Phi^{BR*,BR*} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1}^{US} \\ \mathbf{X}_{t-1}^{BR*} \end{bmatrix} + \begin{bmatrix} \Sigma^{US} & \mathbf{0} \\ \Sigma^{BR*,US} & \Sigma^{BR*,BR*} \end{bmatrix} \varepsilon_t^{BR}.$$

Note that the pricing kernel for the Brazilian defaultable bond returns is $M_{t+1} \Theta_{t+1}$, which is the stochastic discount factor M_{t+1} of the representative buyer of the U.S. non-defaultable bond market adjusted for the default and liquidity risks Θ_{t+1} . Once Θ_{t+1} carries information on the recovery rate in case of default, probability of default and liquidity effects on the price of the bond, we should think of Θ_{t+1} as maturity-specific. So, we should rewrite equation (B-2) as:

$$V_t^{N+1,BRUS} = \mathbb{E}[M_{t+1} \Theta_{t+1}^N V_{t+1}^{N,BRUS}], \quad (2-8)$$

To take the maturity dependence into account, also assume that $\ln \theta_t^N$ has three additive components: a constant, a component linear in the maturity and a component linear on the state variables, so that:

$$\ln \theta_t^N = \theta_0^N + \theta_1 \mathbf{X}_t = N\theta_{00} + \theta_{01} + \theta_1' \mathbf{X}_t. \quad (2-9)$$

A crucial implication of this hypothesis is that θ_t^N varies (linearly) with the maturity, but the effect of the \mathbf{X}_t on θ_t^N does not depend on maturity.

In a model with the characteristics described above, we can write

$$v_t^{N+1,BRUS} = v_t^{N+1,US} + D_{N+1}^{BRUS} + E_{N+1}^{BRUS'} X_t^{BR}, \quad (2-10)$$

where lowercase letters denote variables in log and:

$$\bar{D}_{N+1}^{BRUS} = \theta_0^N + \bar{D}_N^{BRUS} + (\theta_1 + \bar{E}_N^{BRUS})' \Sigma^{BR} [J' (\Sigma^{US'} \bar{B}_N - \lambda_0) \quad (2-11)$$

$$+ \frac{1}{2} \Sigma^{BR'} (\theta_1 + \bar{E}_N^{BRUS})] \quad (2-12)$$

$$\bar{E}_{N+1}^{BRUS'} = (\theta_1 + \bar{E}_N^{BRUS})' (\Phi^{BR} - \Sigma^{BR} J' \lambda_1' J). \quad (2-13)$$

Here, J is a selection matrix such that $X_t^{US} = J X_t^{BRUS}$. Call $y_t^{N,US}$ the yield on the N -period sovereign bond issued by the U.S. government, and $y_t^{N,BRUS}$ the analog for the Brazilian government (USD-denominated). Therefore, $y_t^{N,BRUS} = -\frac{v_t^{N+1,BRUS}}{N} = y_t^{N,US} + D_N^{BRUS} + E_N^{BRUS'} X_t^{BR}$. The spreads $s_t^{N,BRUS} = y_t^{N,BRUS} -$

$y_t^{N,US}$ can be written as

$$s_t^{N,BRUS} = D_N^{BRUS} + E_N^{BRUS'} X_t^{BR},$$

where $D_N^{BRUS} = -\overline{D}_N^{BRUS}/N$ and $E_N^{BRUS} = -\overline{E}_N^{BRUS}/N$. Proof in Appendix G.

Spreads are affine functions of factors. Once we have yields on bonds of different maturities, we estimate the parameters by imposing no-arbitrage restrictions on cross-section estimates given by equations 2-11 and 2-13. So far, we need to estimate the parameters $\Psi^{US} = (\Phi^{US}, \Sigma^{US}, \lambda_0, \lambda_1)$ and $\Psi^{BRUS} = (\Phi^{BR*,US}, \Phi^{BR*,BR*}, \Sigma^{BR*,US}, \Sigma^{BR*,BR*}, \theta_{00}, \theta_{01}, \theta_1)$.

We now turn to the model for the Brazilian bonds denominated in BRL. Our starting point is Proposition 1 in Backus, Foresi and Telmer (2001), that links the pricing kernels denominated in two different currencies. It says that the ratio of the pricing kernel denominated in the foreign currency (“pounds”) to the kernel denominated in the domestic currency (“USD”) equals the depreciation rate of the USD vis-a-vis the pound: If markets are complete, these pricing kernels are unique and so is the ratio of the two. We extrapolate this proposition for the Brazilian bond market where bonds in both foreign and domestic currencies are traded. According to this proposition,

$$\frac{\Lambda_{t+1}^*}{\Lambda_{t+1}} = \frac{S_{t+1}}{S_t},$$

where Λ^* and Λ are, respectively, the pricing kernel denominated in the foreign and domestic currencies and $\frac{S_{t+1}}{S_t}$ is the depreciation rate. From equation (B-2), we know that $\Lambda_{t+1}^* = M_{t+1}\Theta_{t+1}$. So, simple substitution gives

$$\Lambda_{t+1} = M_{t+1}\Theta_{t+1} \frac{S_t}{S_{t+1}}.$$

We can then write the pricing equation for bonds denominated in BRL (that henceforward will receive the superscript $BRRS$) as:

$$V_t^{N+1,BRRS} = \mathbb{E}[M_{t+1}\Theta_{t+1}^N \Pi_{t+1} V_{t+1}^{N,BRRS}], \quad (2-14)$$

where $\Pi_{t+1} = \frac{S_t}{S_{t+1}}$. Following the linear approach, we assume that

$$\log \Pi_t := \pi_t = \pi_0 + \pi_1 \mathbf{X}_t^{BR} \quad (2-15)$$

Equation (2-14) guarantees no-arbitrage. In a model with the characteristics described above, we can write

$$v_t^{N+1,BRRS} = v_t^{N+1,US} + D_{N+1}^{BRUS} + E_{N+1}^{BRUS'} X_t^{BR}, \quad (2-16)$$

where lowercase letters denote variables in log, and:

$$\begin{aligned} \overline{D}_{N+1}^{BRUS} &= \pi_0 + \theta_0^N + \overline{D}_N^{BRRS} + \\ &+ (\pi_1 + \theta_1 + \overline{E}_N^{BRRS})' \Sigma^{BR} [J'(\Sigma^{US'} \overline{B}_N - \lambda_0) + \frac{1}{2} \Sigma^{BR'} (\pi_1 + \theta_1 + \overline{E}_N^{BRRS})] \end{aligned} \quad (2-17)$$

and

$$\overline{E}_{N+1}^{BRUS'} = (\pi_1 + \theta_1 + \overline{E}_N^{BRRS})' (\Phi^{BR} - \Sigma^{BR} J' \lambda_1' J). \quad (2-18)$$

If $y_t^{N,BRRS}$ is the yield on the N -period sovereign bond denominated in BRL issued by the Brazilian government, then $y_t^{N,BRRS} = -\frac{v_t^{N+1,BRRS}}{N} = y_t^{N,US} + D_N^{BRRS} + E_N^{BRRS'} X_t^{BR}$. The spreads $s_t^{N,BRRS} = y_t^{N,BRRS} - y_t^{N,US}$ can be written as

$$s_t^{N,BRRS} = D_N^{BRRS} + E_N^{BRRS'} X_t^{BR},$$

where $D_N^{BRRS} = -\overline{D}_N^{BRRS}/N$ and $E_N^{BRRS} = -\overline{E}_N^{BRRS}/N$. Proof in Appendix B.

The parameters that are left to estimate in this step are only $\Psi^{BRRS} = (\pi_0, \pi_1)$. π_0 is a scalar, and π_1 is a K -dimensional vector.

We follow Ang and Piazzesi (2003) and estimate the model in a multi-step procedure. In the first step, we estimate the parameters of the dynamics of U.S. factors in equation (2-7) obtaining $\hat{\Phi}^{US}$ and $\hat{\Sigma}^{US}$. In a second step, we estimate the equation (2-5) under restrictions in equations (F-1) and (F-2) and obtain the parameters $\hat{\lambda}_0$ and $\hat{\lambda}_1$. We assume the simplifying hypothesis that the matrix λ_1 is diagonal. The first two steps yields $\hat{\Psi}^{US}$. In the third step, we estimate the parameters of the dynamics of Brazilian factors in equation (2-7) obtaining $\hat{\Phi}^{BR*,US}$, $\hat{\Phi}^{BR*,BR*}$ and $\hat{\Sigma}^{BR*,US}$, $\hat{\Sigma}^{BR*,BR*}$. In the fourth step, we estimate the equation (2-10) and obtain $\hat{\theta}_0^N$ and $\hat{\theta}_1$. These two last steps give us $\hat{\Psi}^{BRUS}$. The fifth and last step yields $\hat{\Psi}^{BRRS}$, $\hat{\pi}_0$ and $\hat{\pi}_1$ through the estimation of equation (2-16).

2.3 Data

We use prices of zero-coupon bonds with maturities of 3, 6, 12, 24, 36, and 60 months, on monthly basis from January 1999 to May 2012, extracted from Bloomberg. Missing values in the Brazilian bond-yield series were filled by cubic spline interpolation. We also used an outlier filter in these series to avoid movements of over 50% within a month. Factors are divided into a risk indicator, macroeconomic variables, yields- and spreads-related variables.

The risk indicator is Moody's Baa Corporate Bond Yield (HY), from the Board of Governors of the Federal Reserve System. For the United States, the macroeconomic variables are measures of activity and inflation. The measures of activity is the first principal component of important indicators: monthly changes

of non-farm payroll and industrial production, unemployment rate and capacity utilization. The measures of inflation is the first principal component of monthly changes of consumer price index, producer price index and commodities prices. Data are seasonally adjusted and are from the BLS. For Brazil, the macroeconomic variables are measures of activity, inflation, external solvency, and deficit ex-interest rates. The measure of activity is the first principal component of monthly changes of GDP (Brazilian Central Bank estimate), industrial production and capacity utilization in the industrial sector and unemployment rate in São Paulo. The measure of inflation is the first principal component of monthly changes of IPCA and IGP-M. Data are seasonally adjusted and are from Ipeadata. External solvency indicator is the first principal component of external debt, real exchange rate, terms of trade, current account flow, openness (imports plus exports as a ratio of GDP), international prices of commodities (CCI) and foreign currency reserves owned by the Central Bank. Data source is Ipeadata, except CCI, that is from Bloomberg.

The yields-related variables are the three first principal components of yields (for the U.S) or spreads (for Brazil) estimated from an eigenvalue decomposition of the observed yields (spreads) covariance matrix, orthogonalized with respect to the macroeconomics and risk factors (Cochrane and Piazzesi [2008]).¹ These yields and spreads factors represent variables other than factors considered above, i.e., the portion of yields and spreads curve not explained by economic fundamentals.

It has been shown by this literature (Litterman and Scheinkman [1992], Ang and Piazzesi [2003]) that up to three latent factors are appropriate to capture most salient features of the yield curve. These principal components can be understood as the portion macroeconomic and risk variables that are unspanned by bond yields, in the terminology of Joslin, Priebsch and Singleton (2010). This way, in the model for the U.S. bond yields, we will be able to write our pricing equation as $v_t^N = \bar{A}_N + \bar{B}_N^M M_t + \bar{B}_N^P P_t$, where M_t represents the macroeconomic variables, P_t , the first three first principal components of yields orthogonalized with respect to the macroeconomics and risk factors. By construction, each of the macroeconomic variables will be almost spanned by v and P . It will not be perfectly spanned because of the information on the principal components of orders four and larger, which are neglectable.² Even if macroeconomic and risk indicators were spanned by the conventional (non-orthogonalized) three latent factors, it would be useful

¹More specifically, the U.S. yields-related variables are the three first principal components of yields estimated from an eigenvalue decomposition of the observed yields covariance matrix, orthogonalized with respect to U.S. measures of activity and inflation and HY . Brazilian spreads-related variables are the three first principal components of yields estimated from an eigenvalue decomposition of the observed yields covariance matrix, orthogonalized with respect to HY , U.S. measures of activity and inflation, U.S. yields-related variables, and Brazilian economy variables described above.

²For the U.S. model, the portion variance of yields not explained by macroeconomic and yields variables variables is as small as 0.3%. For the case of spreads, this portion is of 2.4%.

for our purpose to isolate them. This way, we can extract from bond yields all the information concerning macroeconomic and risk factors. As in Joslin, Pribsch and Singleton (2010), we find that a good deal of information on the yield curve is in the portion unspanned by macroeconomic variables, but in this paper, we are interested in understanding the role of the spanned portion. As a practical matter, these remaining factors have the role of capturing most of the persistence of yields (and spreads, for Brazil).

Industrial production and inflation were included in the light of the findings of Ang and Piazzesi (2003), which shows the importance of these variables, along with principal components of yields, in their term structure model applied to the US data. External solvency, deficit, internal debt were included to capture the fiscal effect on spreads. Moody's Baa Corporate Bond Yield was chosen to capture international finance market stress.

2.4 Results

2.4.1 What has been driving spreads?

Tables 2.2 and 2.1 show parameter estimates for the U.S. non-defaultable and the Brazilian defaultable bonds model, respectively. In both models, standard errors are calculated through a non-parametric Bootstrap procedure (400 repetitions).³

The reported standard errors for the defaultable bonds model parameters $\hat{\Psi}^{BRUS}$ and $\hat{\Psi}^{BRRS}$ estimated in the third to fifth steps described in the last paragraph of section 2.2 take $\hat{\Psi}^{US}$ as constant. It means that through the 400 repetitions of steps three to five described in the end of section 2.2, we keep the estimated parameters in $\hat{\Psi}^{US}$ frozen. Besides becoming estimation easier, this method is in line with our approach to consider the foreign leader economy dynamics as exogenous to the Brazilian economy.

³See Horowitz (2001) and Davison and Hinkley (1997).

Table 2.1: Parameter Estimates - Model for the U.S. Term Structure of Yields

VAR of Factors X_t							
$X_t^{BR} = \Phi^{BR} X_{t-1}^{BR} + \Sigma^{BR} \varepsilon_t^{BR}$							
Matrix Φ^{US}							
	HY _{t-1}	Activity ^{US} _{t-1}	Inflation ^{US} _{t-1}	Princ. Comp1 ^{US} _{t-1}	Princ. Comp2 ^{US} _{t-1}	Princ. Comp3 ^{US} _{t-1}	
HY _t	0.918 (0.165)***	0.027 (0.018)	0.042 (0.023)*	0.026 (0.019)	0.014 (0.019)	-0.034 (0.019)*	
Activity ^{US} _t	-0.037 (0.03)	0.832 (0.161)***	0.049 (0.036)	0.109 (0.029)***	0.074 (0.028)**	0.030 (0.03)	
Inflation ^{US} _t	-0.103 (0.075)	0.120 (0.08)	0.206 (0.098)**	0.048 (0.076)	0.037 (0.075)	0.121 (0.08)	
Princ. Comp1 ^{US} _t	0.037 (0.05)	0.122 (0.05)**	-0.023 (0.048)	0.715 (0.159)***	-0.079 (0.047)*	0.694 (0.05)	
Princ. Comp2 ^{US} _t	0.024 (0.047)	-0.039 (0.046)	0.105 (0.046)**	-0.065 (0.048)	0.738 (0.16)***	0.028 (0.047)	
Princ. Comp3 ^{US} _t	0.101 (0.05)**	-0.079 (0.049)	0.193 (0.053)***	0.016 (0.045)	-0.065 (0.048)	0.694 (0.157)***	
Matrix Σ^{US}							
	HY	Activity ^{US}	Inflation ^{US}	Princ. Comp1 ^{US}	Princ. Comp2 ^{US}	Princ. Comp3 ^{US}	
HY	0.267 (0.123)**	-	-	-	-	-	
Activity ^{US}	-0.001 (0.04)	0.386 (0.093)***	-	-	-	-	
Inflation ^{US}	-0.210 (0.12)*	0.060 (0.083)	0.901 (0.068)***	-	-	-	
Princ. Comp1 ^{US}	-	-	-	0.623 (0.081)***	-	-	
Princ. Comp2 ^{US}	-	-	-	0.196 (0.047)***	0.563 (0.072)***	-	
Princ. Comp3 ^{US}	-	-	-	0.063 (0.052)	0.013 (0.056)	0.631 (0.071)***	
Short Rate Regression							
$spread_t^{US} = \delta_0^{US} + \delta_1^{US} X_t^{US} + u_t^{BR}$							
	100* δ_0	Vector 100* δ_1					
	2.358 (0.002)***	HY _t 0.930 (0.002)***	Activity ^{US} _t 1.482 (0.002)***	Inflation ^{US} _t 0.095 (0.002)***	Princ. Comp1 ^{US} _t 1.024 (0.002)***	Princ. Comp2 ^{US} _t -0.208 (0.002)***	Princ. Comp3 ^{US} _t 0.087 (0.002)***
Market Price of Risk							
$\lambda_t^{US} = \lambda_0^{US} + \lambda_1^{US} X_t^{US}$							
	HY _t	Activity ^{US} _t	Inflation ^{US} _t	Princ. Comp1 ^{US} _t	Princ. Comp2 ^{US} _t	Princ. Comp3 ^{US} _t	
	-0.110 (9.117)	-0.393 (5.968)	-2.651 (7.37)	0.620 (4.571)	0.068 (2.43)	0.121 (4.105)	
Diagonal of Matrix λ_1							
	HY _t	Activity ^{US} _t	Inflation ^{US} _t	Princ. Comp1 ^{US} _t	Princ. Comp2 ^{US} _t	Princ. Comp3 ^{US} _t	
	-0.131 (0.858)	0.052 (0.062)	0.338 (0.39)	-0.018 (0.052)	-0.141 (0.182)	2.390 (0.685)***	

Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.2: Parameter Estimates - Model for the Brazilian Term Structure of Spreads

VAR of factors X_t		Matrix $(\beta^{100\%}, \beta^{100\%}, \beta^{100\%})$											
		HY_t	$Activity^{100\%}_t$	$Inflation^{100\%}_t$	$Princ. Comp1^{100\%}_t$	$Princ. Comp2^{100\%}_t$	$Princ. Comp3^{100\%}_t$	$Princ. Comp4^{100\%}_t$	$Princ. Comp5^{100\%}_t$	$Princ. Comp6^{100\%}_t$	$Princ. Comp7^{100\%}_t$	$Princ. Comp8^{100\%}_t$	$Princ. Comp9^{100\%}_t$
$Activity^{100\%}_t$	0.010 (0.097)	0.093 (0.081)	0.324 (0.079)***	0.055 (0.079)***	0.137 (0.079)***	0.159 (0.079)***	-0.096 (0.079)***	-0.088 (0.079)***	-0.118 (0.079)***	0.137 (0.079)***	0.060 (0.079)***	-0.007 (0.079)***	0.075 (0.079)***
$Inflation^{100\%}_t$	0.065 (0.081)	-0.051 (0.079)***	0.139 (0.079)***	-0.125 (0.079)***	0.051 (0.079)***	0.020 (0.079)***	0.014 (0.079)***	0.640 (0.079)***	-0.103 (0.079)***	0.214 (0.079)***	0.154 (0.079)***	0.059 (0.079)***	0.075 (0.079)***
$External Solvency^{100\%}_t$	0.081 (0.097)	0.051 (0.097)	0.027 (0.097)	0.044 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)	0.040 (0.097)
$Deficit^{100\%}_t$	0.177 (0.097)	-0.165 (0.097)	-0.147 (0.097)	-0.041 (0.097)	-0.022 (0.097)	-0.087 (0.097)	-0.086 (0.097)	-0.027 (0.097)	0.147 (0.097)	0.039 (0.097)	0.062 (0.097)	-0.115 (0.097)	-0.079 (0.097)
$Princ. Comp1^{100\%}_t$	-0.021 (0.052)	0.066 (0.046)	-0.078 (0.046)	-0.032 (0.046)	0.047 (0.046)	0.034 (0.046)	-0.100 (0.046)	0.064 (0.046)	0.069 (0.046)	0.058 (0.046)	0.030 (0.046)	0.065 (0.046)	-0.082 (0.046)
$Princ. Comp2^{100\%}_t$	0.051 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)	0.041 (0.051)
$Princ. Comp3^{100\%}_t$	0.060 (0.077)	0.071 (0.059)	0.055 (0.059)	-0.087 (0.059)	0.032 (0.059)	-0.071 (0.059)	0.041 (0.059)	0.041 (0.059)	0.041 (0.059)	0.041 (0.059)	0.041 (0.059)	0.041 (0.059)	0.041 (0.059)
$Princ. Comp4^{100\%}_t$	-0.002 (0.069)	0.007 (0.069)	0.001 (0.069)	-0.006 (0.069)	0.006 (0.069)	0.001 (0.069)	0.875 (0.069)	-	-	-	-	-	-
$Princ. Comp5^{100\%}_t$	-0.003 (0.044)	0.001 (0.044)	-0.001 (0.044)	0.002 (0.044)	0.002 (0.044)	0.001 (0.044)	0.029 (0.044)	0.536 (0.044)	-	-	-	-	-
$Princ. Comp6^{100\%}_t$	0.011 (0.087)	0.007 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)	0.006 (0.087)
$Princ. Comp7^{100\%}_t$	-0.005 (0.068)	0.005 (0.068)	0.006 (0.068)	-0.007 (0.068)	-0.012 (0.068)	0.001 (0.068)	-0.064 (0.068)	-0.007 (0.068)	0.860 (0.068)	-	-	-	-
$Princ. Comp8^{100\%}_t$	-	-	-	-	-	-	-	-	-	0.494 (0.053)***	-	-	-
$Princ. Comp9^{100\%}_t$	-	-	-	-	-	-	-	-	-	0.452 (0.059)***	-	-	-
$Princ. Comp10^{100\%}_t$	-	-	-	-	-	-	-	-	-	0.279 (0.067)***	0.019 (0.055)***	-	-
$Princ. Comp11^{100\%}_t$	-	-	-	-	-	-	-	-	-	-	0.688 (0.055)***	-	-

Short Rate Regression - Bonds Denominated in 100% ξ_t		Matrix $(\beta^{100\%}, \beta^{100\%}, \beta^{100\%})$											
		HY_t	$Activity^{100\%}_t$	$Inflation^{100\%}_t$	$Princ. Comp1^{100\%}_t$	$Princ. Comp2^{100\%}_t$	$Princ. Comp3^{100\%}_t$	$Princ. Comp4^{100\%}_t$	$Princ. Comp5^{100\%}_t$	$Princ. Comp6^{100\%}_t$	$Princ. Comp7^{100\%}_t$	$Princ. Comp8^{100\%}_t$	$Princ. Comp9^{100\%}_t$
$spread^{100\%}_t = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	2.254 (0.077)***	0.539 (0.022)***	-0.725 (0.019)***	-0.155 (0.019)***	-0.413 (0.018)***	-0.131 (0.018)***	0.074 (0.018)***	1.041 (0.018)***	1.041 (0.018)***	1.041 (0.018)***	1.041 (0.018)***	1.041 (0.018)***	1.041 (0.018)***
$spread^{100\%}_t = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	12.482 (0.025)***	-0.835 (0.025)***	-0.569 (0.025)***	-0.488 (0.025)***	-1.549 (0.025)***	0.042 (0.025)***	-0.179 (0.025)***	0.087 (0.025)***	0.887 (0.025)***	0.376 (0.025)***	2.743 (0.025)***	-1.355 (0.025)***	-0.535 (0.025)***

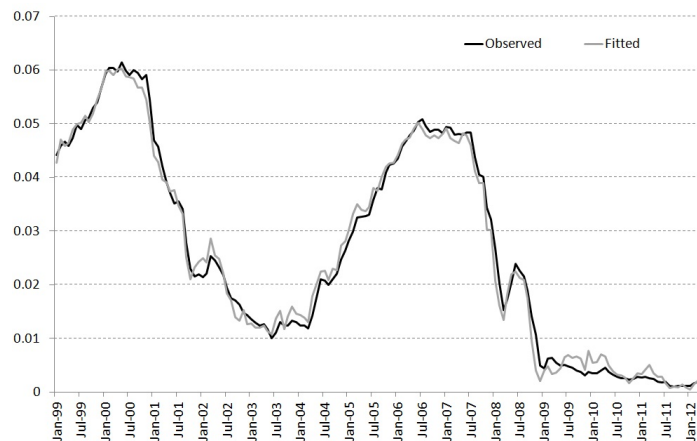
Short Rate Regression - Bonds Denominated in Reals		Matrix $(\beta^{100\%}, \beta^{100\%}, \beta^{100\%})$											
		HY_t	$Activity^{100\%}_t$	$Inflation^{100\%}_t$	$Princ. Comp1^{100\%}_t$	$Princ. Comp2^{100\%}_t$	$Princ. Comp3^{100\%}_t$	$Princ. Comp4^{100\%}_t$	$Princ. Comp5^{100\%}_t$	$Princ. Comp6^{100\%}_t$	$Princ. Comp7^{100\%}_t$	$Princ. Comp8^{100\%}_t$	$Princ. Comp9^{100\%}_t$
$spread^{100\%}_t = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	0.002 (0.062)	-0.005 (0.064)	0.021 (0.013)**	0.027 (0.013)**	-0.015 (0.012)	-0.006 (0.012)	0.014 (0.012)	-0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)
$spread^{100\%}_t = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	0.002 (0.062)	-0.005 (0.064)	0.021 (0.013)**	0.027 (0.013)**	-0.015 (0.012)	-0.006 (0.012)	0.014 (0.012)	-0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)

Recovery Intensity		Matrix $(\beta^{100\%}, \beta^{100\%}, \beta^{100\%})$											
		HY_t	$Activity^{100\%}_t$	$Inflation^{100\%}_t$	$Princ. Comp1^{100\%}_t$	$Princ. Comp2^{100\%}_t$	$Princ. Comp3^{100\%}_t$	$Princ. Comp4^{100\%}_t$	$Princ. Comp5^{100\%}_t$	$Princ. Comp6^{100\%}_t$	$Princ. Comp7^{100\%}_t$	$Princ. Comp8^{100\%}_t$	$Princ. Comp9^{100\%}_t$
$\ln(\Gamma^t) = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	0.002 (0.062)	-0.005 (0.064)	0.021 (0.013)**	0.027 (0.013)**	-0.015 (0.012)	-0.006 (0.012)	0.014 (0.012)	-0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)
$\ln(\Gamma^t) = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	0.002 (0.062)	-0.005 (0.064)	0.021 (0.013)**	0.027 (0.013)**	-0.015 (0.012)	-0.006 (0.012)	0.014 (0.012)	-0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)	0.012 (0.012)

Exchange Rate		Matrix $(\beta^{100\%}, \beta^{100\%}, \beta^{100\%})$											
		HY_t	$Activity^{100\%}_t$	$Inflation^{100\%}_t$	$Princ. Comp1^{100\%}_t$	$Princ. Comp2^{100\%}_t$	$Princ. Comp3^{100\%}_t$	$Princ. Comp4^{100\%}_t$	$Princ. Comp5^{100\%}_t$	$Princ. Comp6^{100\%}_t$	$Princ. Comp7^{100\%}_t$	$Princ. Comp8^{100\%}_t$	$Princ. Comp9^{100\%}_t$
$\ln(\Gamma^t) = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	-0.517 (0.074)	-0.038 (0.074)	0.017 (0.074)	0.080 (0.074)	-0.065 (0.074)	-0.037 (0.074)	0.011 (0.074)	0.007 (0.074)	0.008 (0.074)	0.007 (0.074)	0.007 (0.074)	0.007 (0.074)	0.007 (0.074)
$\ln(\Gamma^t) = \theta^{100\%}_0 + \theta^{100\%}_1 r_t + \theta^{100\%}_2 \pi_t + \theta^{100\%}_3 s_t + \theta^{100\%}_4 \delta_t$	-0.517 (0.074)	-0.038 (0.074)	0.017 (0.074)	0.080 (0.074)	-0.065 (0.074)	-0.037 (0.074)	0.011 (0.074)	0.007 (0.074)	0.008 (0.074)	0.007 (0.074)	0.007 (0.074)	0.007 (0.074)	0.007 (0.074)

Standard errors in parentheses. *, significant at 10%; **, significant at 5%; ***, significant at 1%.

Figure 2.3: U.S. 1-Year Bond-Yield: Observed and Fitted Values



Note: A yield of $x\%$ per year enters as $z = \log(1+x/100)$.

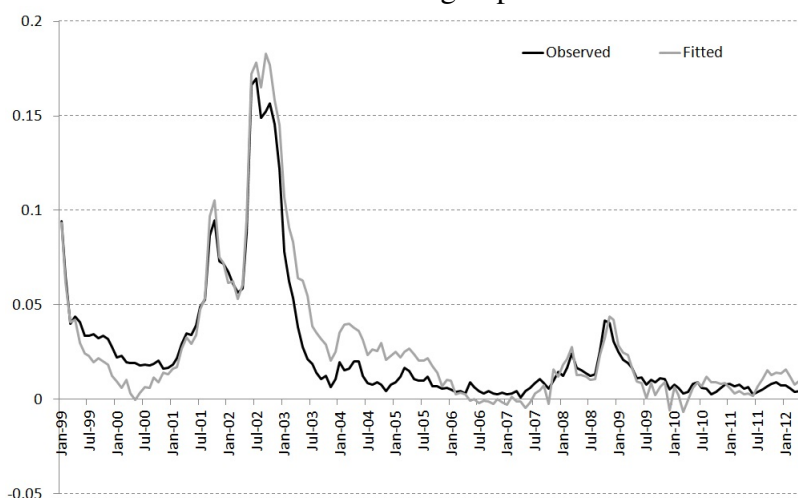
The fit of our model for the U.S. bond yields and Brazilian sovereign spreads are shown in figures 2.3 to 2.5. The no-arbitrage model seems to have a good fit, capturing the major salient features of the observed series. It is clear that the fit for the U.S. curve is better, which is a consequence of our modelling option: in the model for the spreads, we take the estimated U.S. bond-yields as the observed ones, adding more noise to the estimation of spreads (this feature of the model is clear in the proof of equations [2-10] to [2-13] in Appendix G and of equations [2-16] to [2-18] in Appendix H).

Recall that we are splitting the discount on the price of a BRL-denominated bond into a portion attributed to default risk and another portion attributed to currency risk. The first portion is estimated in the model for the Brazilian USD-denominated bonds, which is possible because of our assumption that Brazilian USD- and BRL-denominated bonds have a common default risk. The second portion is estimated in the model for the Brazilian USD-denominated bonds.

Note that that the smaller the recovery intensity Θ , the larger the discount on the price of the defaultable USD bond due to default risk. The estimated coefficients on Θ shown in table 2.1 indicate that this discount varies negatively with Brazilian inflation and positively with HY , U.S. activity and the Brazilian external solvency. The coefficient on Brazilian activity also positive, even though it is not statistically significant. Note also that the coefficient on maturity is not statistically significant, which suggest either that our specification is deeply mistaken or the recovery intensity does not significantly varies with maturity.

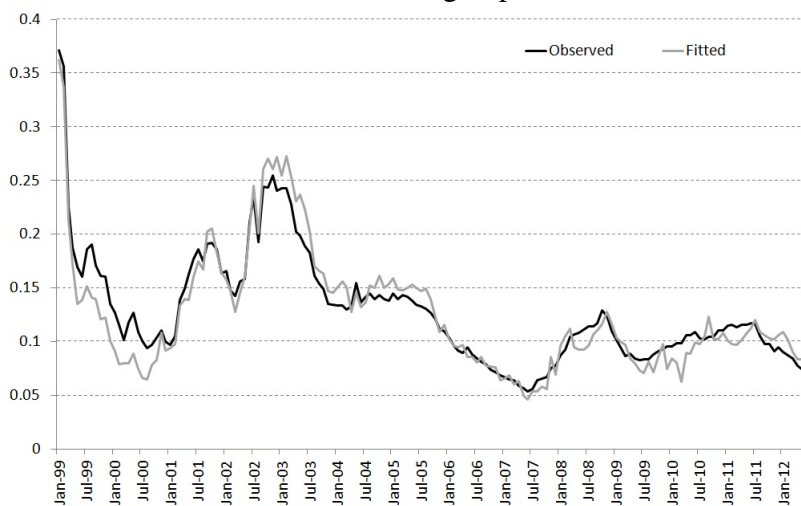
Figure 2.6 shows the estimated path of Θ for one year maturity bond. There are at some remarkable features in this path. First, Θ reached its lowest levels in the beginning of 1999 and in the end of 2002. These two moments were indeed two of the most critical ones for the Brazilian economy within the years in sample. In 1999,

Figure 2.4: 1-Year USD Bond-Yield Sovereign Spreads: Observed and Fitted Values

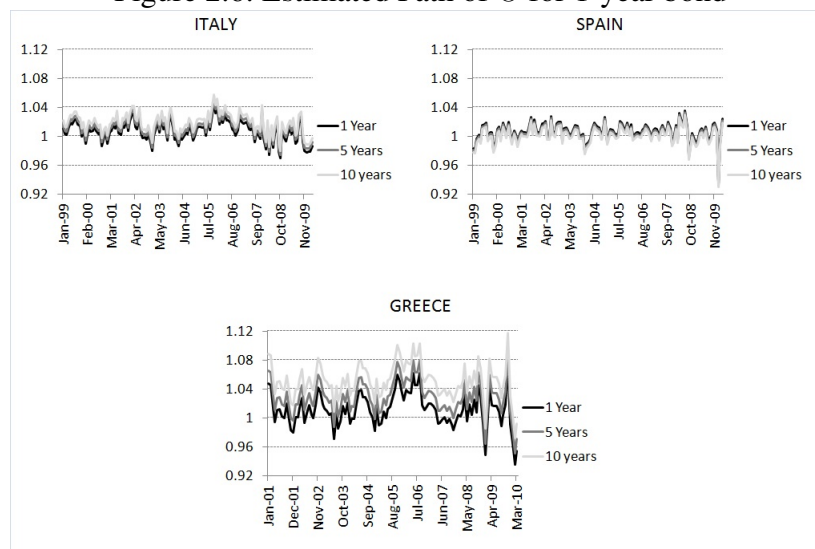


Note: A yield of $x\%$ per year enters as $z = \log(1+x/100)$. If the yield on the Brazilian bond is $xbr\%$ per year and the yield on the U.S. bond is $xus\%$ per year, the spread enters as $s=zbr-zus$, where $zbr = \log(1+xbr/100)$ and $zus = \log(1+xus/100)$.

Figure 2.5: 1-Year BRL Bond-Yield Sovereign Spreads: Observed and Fitted Values



Note: A yield of $x\%$ per year enters as $z = \log(1+x/100)$. If the yield on the Brazilian bond is $xbr\%$ per year and the yield on the U.S. bond is $xus\%$ per year, the spread enters as $s=zbr-zus$, where $zbr = \log(1+xbr/100)$ and $zus = \log(1+xus/100)$.

Figure 2.6: Estimated Path of Θ for 1-year bond

BRL suffered a maxi-devaluation, and fluctuated wildly, after a period since 1994 of a quasi-fixed exchange rate. In the end of 2002, the former president Lula was about to be elected, fact that caused a huge confidence crisis. We can also note in the path of Θ some other smoother movements. The low peak in Θ in 2005 (around and September) its high peak in 2008 (around October) mach moments of elevation (2005) and reduction (2008) of U.S. Fed Funds (see figure 2.7).⁴

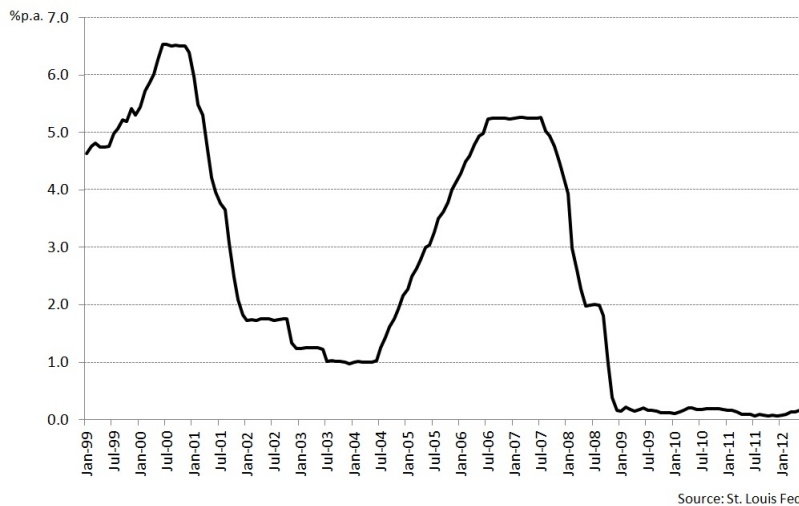
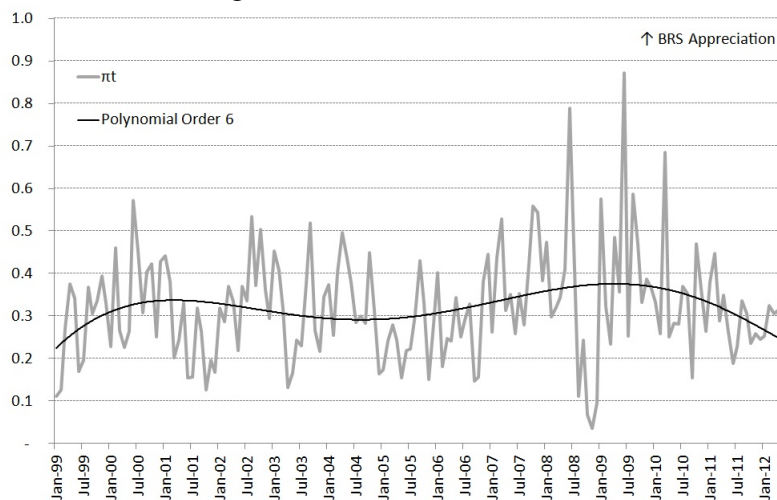
The coefficients of the exchange rate appreciation equation all statistically non-significant. This result may mean that the expected discount on the price of the defaultable BRL bond due specifically to exchange rate risk does not vary with economic fundamentals. Figure 2.8 shows the estimated path of Π and its trend line (a polynomial of order 6). It seems really hard to associate this path to economic facts.

2.4.2 Factor loadings along the spread curve and spread composition

The effect of each factor on the spreads curve is determined by the weights E_N that the term structure model assigns on each spread of maturity N . Figure 2.9 plots the weights of HY and factors related to the Brazilian economy - activity, inflation, external solvency and deficit - on USD and BRL bond-yield spreads over maturities. For both measures of spreads, HY enters negatively for shorter maturities and cause a positive increasing effect for larger maturities. Even though the effect on short maturities is hard to explain, the positive inclination of the effect of HY on maturities is line with what is observed in markets in moments of crisis.

⁴Arora and Cerisola (2001) present evidence that the level of U.S. interest rates has direct positive effects on emerging countries's sovereign bond-yield spreads.

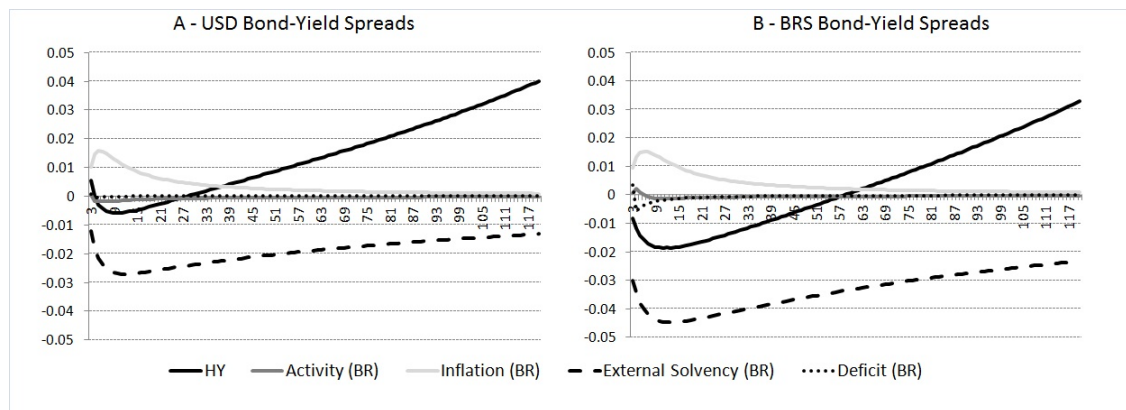
Figure 2.7: Effective Fed Funds Rate

Figure 2.8: Estimated Path of Π 

For shorter maturities, the variable that causes the largest impact is external solvency. The fact that weights E_N are similar for both USD and BRL bond-yield spreads not only in the sign but also in magnitude is due to the fact that we were not able to properly model Π_t as a function of economic fundamentals (see figure 2.8 and table 2.2)).

Figures 2.10 and 2.11 show the estimated composition of spreads on 1-year USD and BRL bonds, respectively. In these figures, we report D_{12} (constant) and the weights E_{12} multiplied by factors over time. Brazilian and U.S. factors (ex- HY) were summed up. Note that until July 2006, Brazilian factors had a positive contribution to spreads and, from then on, their contribution became negative. The results also suggest some other intuitive features of the path of spreads. The first two high peaks in the beginning of 1999 and in the end of 2002 could be explained

Figure 2.9: Weights E_N on USD and BRL bond-yield spreads over maturities N . Weights are multiplied by 100.



mainly by Brazilian factors, and the high peak in end of 2008 had its origin in the U.S. economy. Besides, Brazilian factors have increased their negative contribution to spreads since 2009, specially in the case of BRL bond-yield spreads.

In figures 2.12 and 2.13 we can see the estimated composition of spreads on 1-year USD and BRL bonds, respectively, opened by the Brazilian factors. Again, in these figures, we report D_{12} (constant) and the weights E_{12} multiplied by factors over time. Now, only U.S. factors (ex-*HY*) are summed up.

Figures 2.12 and 2.13 indicate that, among Brazilian factors, external solvency is the most relevant individual factor in the composition of both 1-year USD and BRL bond-yield spreads since the end of 2007. This result had already been suggested by loadings on figure 2.9, that also suggested that inflation would be the second most relevant Brazilian factors for maturities around 1 year. The 1999 high peak for both U.S. and BRL bond-yield spreads could not be explained by economic fundamentals (it was driven mainly by Brazilian principal components of spreads - *BRA PC*), when BRL suffered an unexpected maxi-devaluation. Economic fundamentals were not able to explain the first moments of U.S. bond-yield spreads high in the second semester of 2002. Note that, in the first moments, the spread were driven by *BRA PC*. Its importance is reduced and, in December 2002, spreads were still high and were explained mainly by external (in)solvency and inflation. It is a case of self-fulfilling prophecy.

Figures 2.14 and 2.15 show the estimated composition of spreads on 1-year USD and BRL bonds, respectively, opened by the U.S. factors. Again, in these figures, we report D_{12} (constant) and the weights E_{12} multiplied by factors over time. Among U.S. factors, we can see the great positive influence of U.S. activity on BRL and USD bond yield spreads, specially in critical moments of the recent financial crisis (2007-2011). The figures suggest that non-fundamental U.S. factors

Figure 2.10: 1-year USD bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by HY and Brazilian and U.S. factors over time.

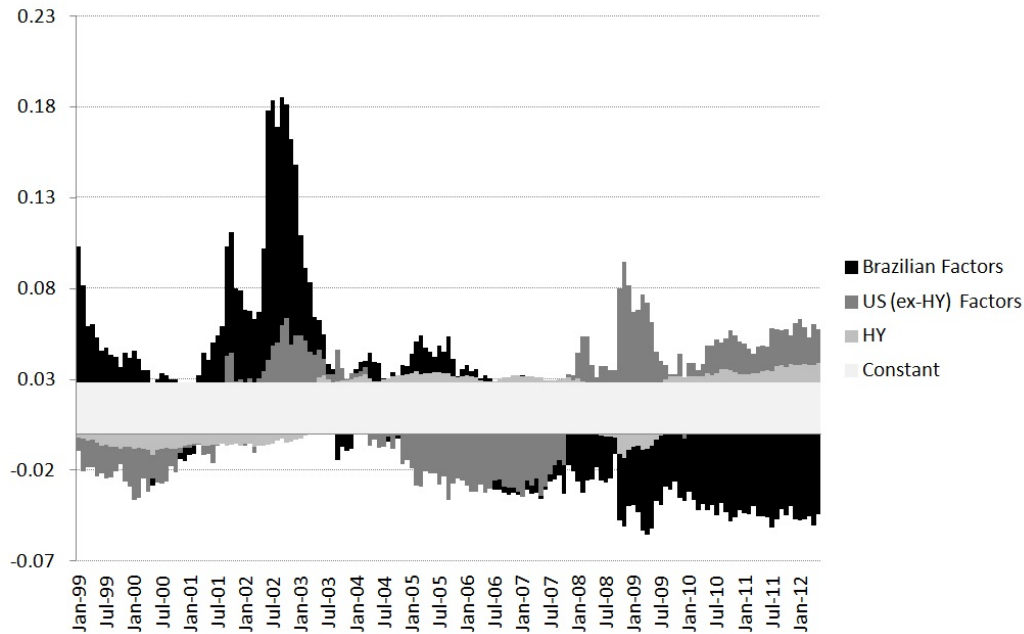


Figure 2.11: 1-year BRL bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by HY and Brazilian and U.S. factors over time.

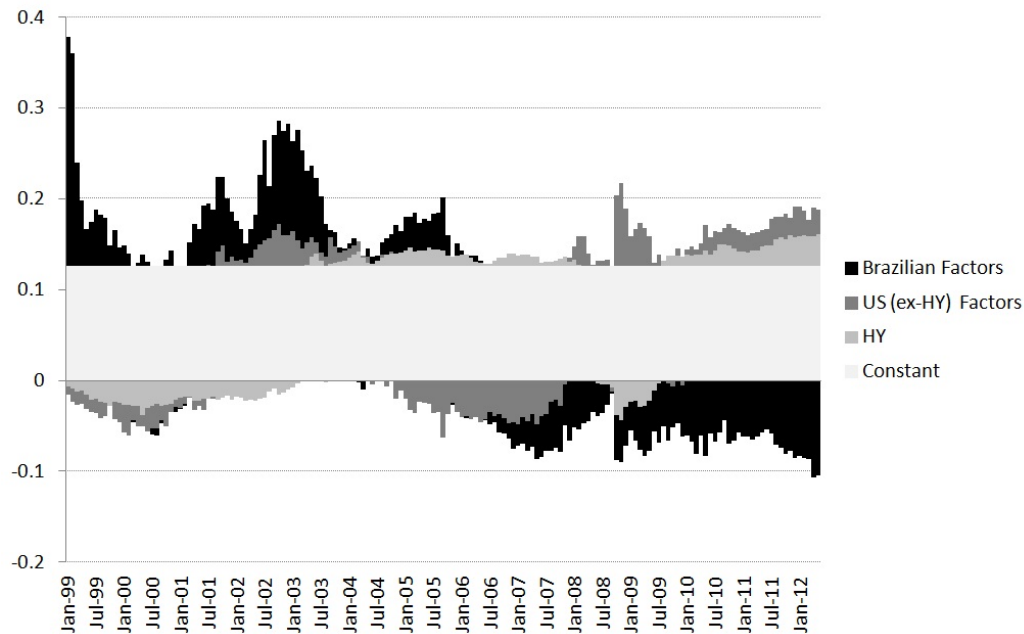


Figure 2.12: 1-year USD bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by factors over time.

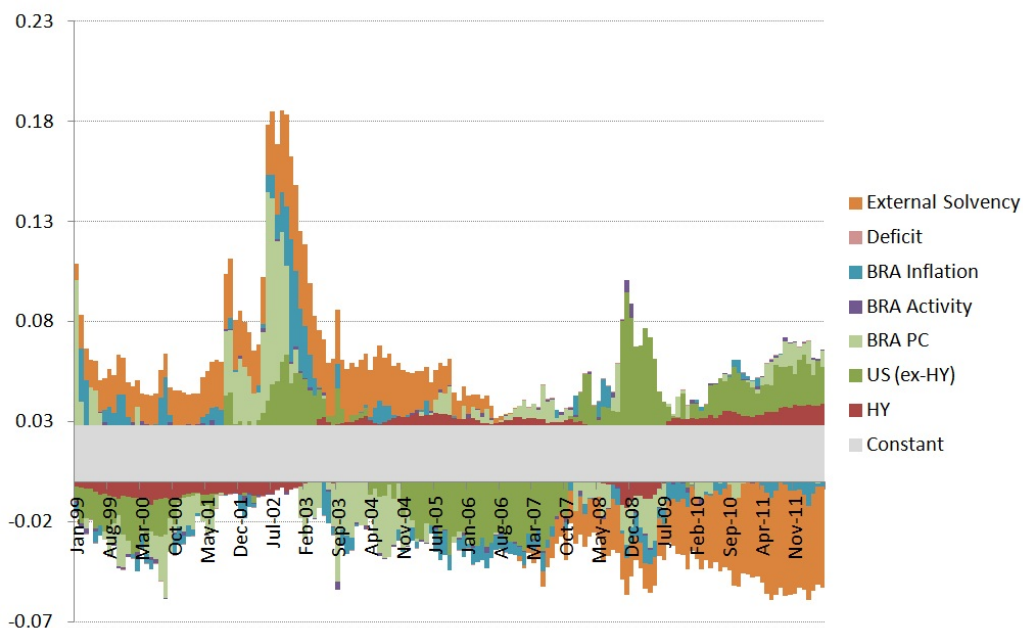


Figure 2.13: 1-year BRL bond-yield spreads: D_{12} (constant) and weights E_{12} multiplied by factors over time.

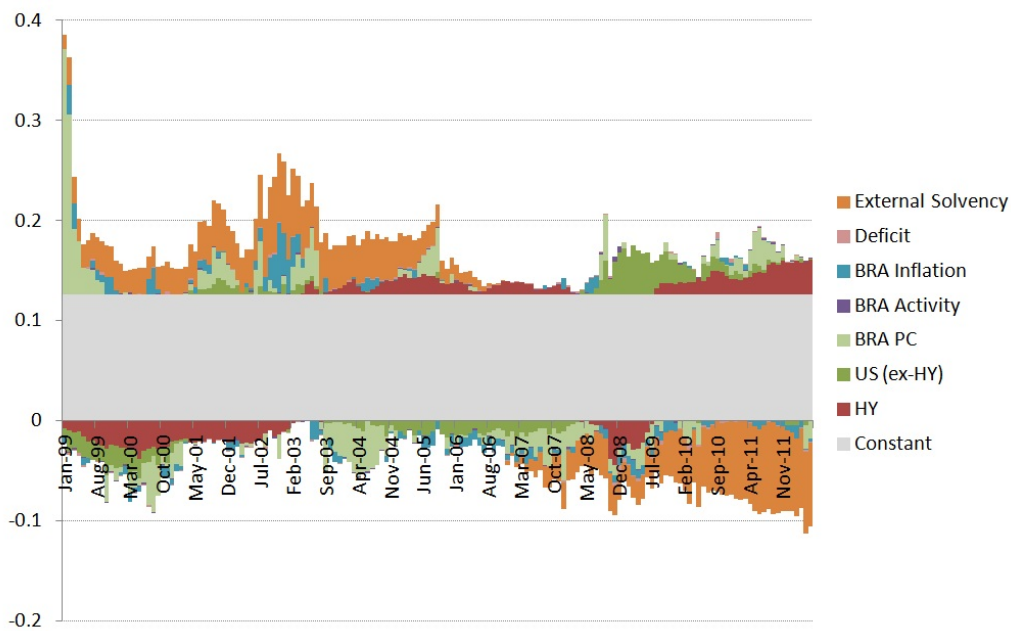


Figure 2.14: 1-year USD bond-yield spreads: Weights E_{12} multiplied by U.S. factors over time.

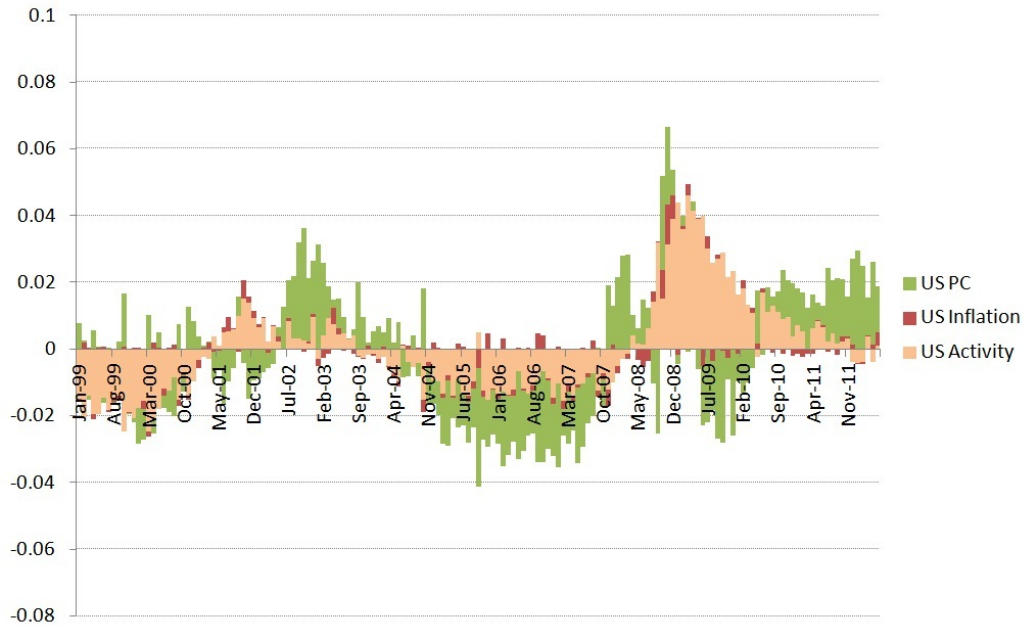


Figure 2.15: 1-year BRL bond-yield spreads: Weights E_{12} multiplied by U.S. factors over time.

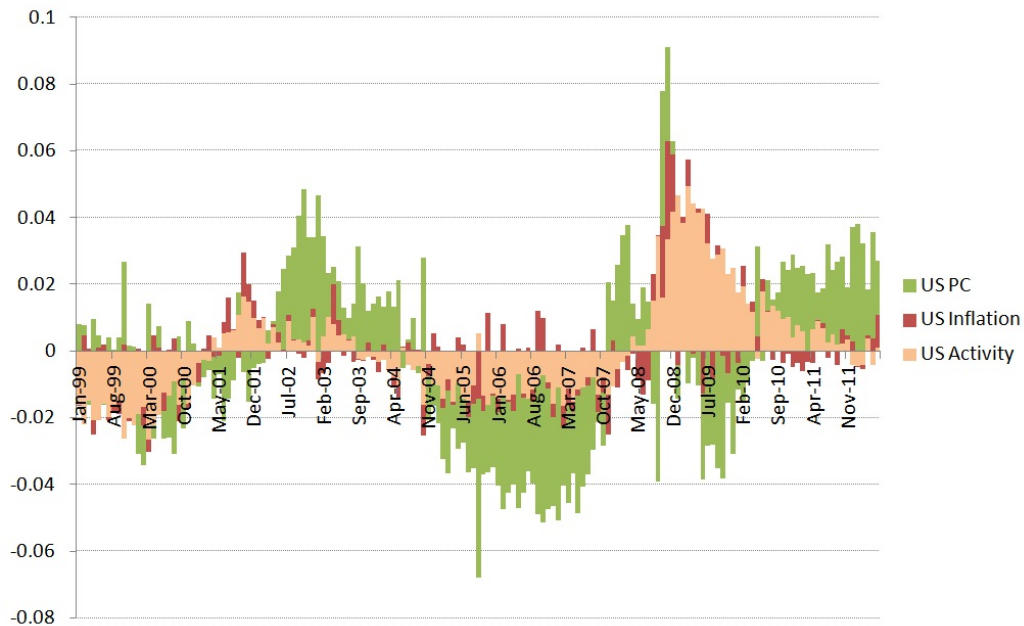
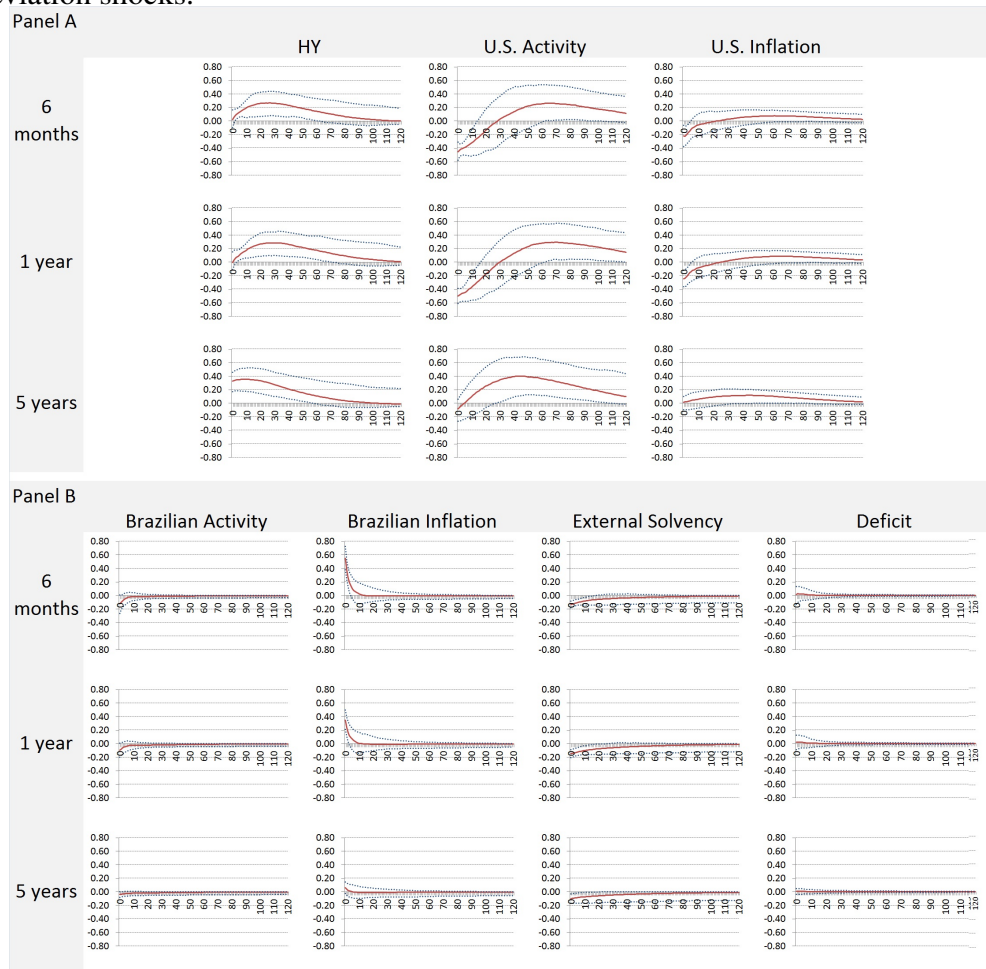


Figure 2.16: USD bond-yield spreads: Impulse Responses from a one standard deviation shocks.



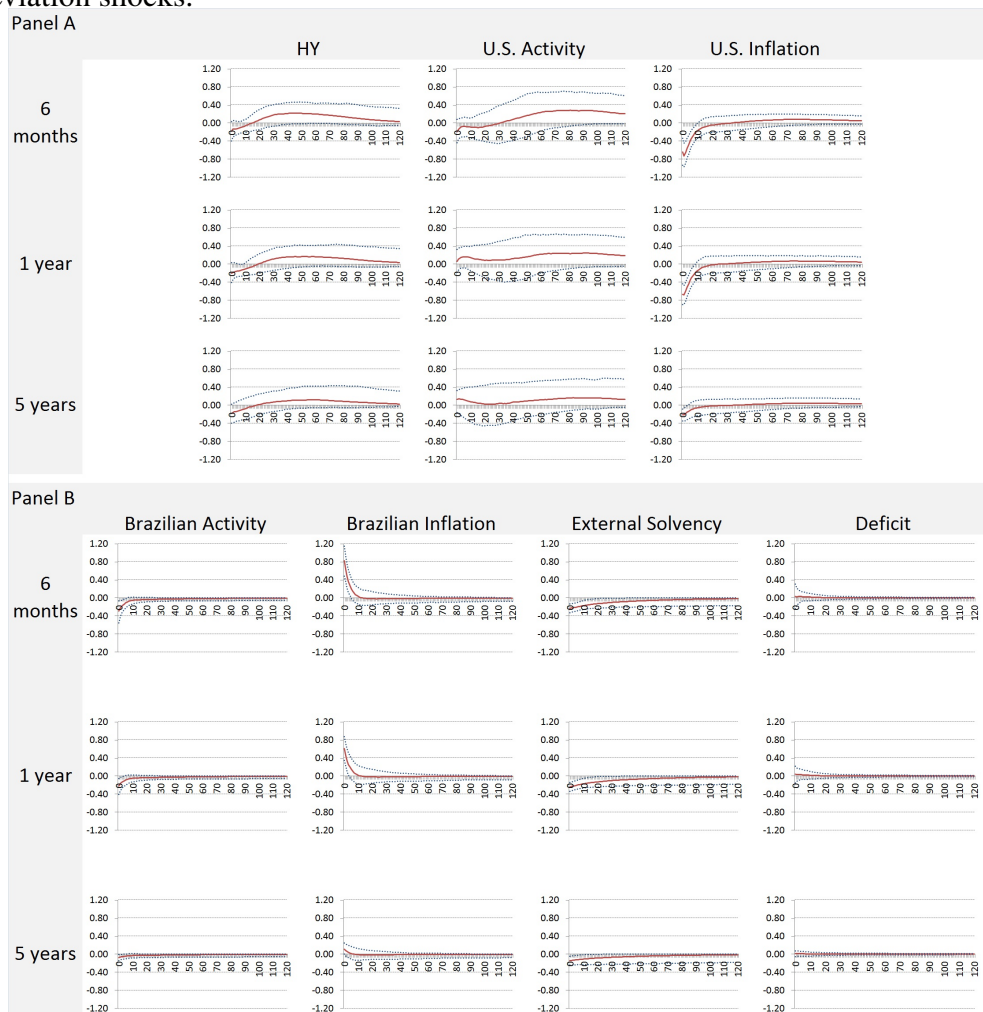
explain the major portion of the positive influence of U.S. factors over spreads, which may be a consequence of the persistence of uncertainty in the aftermath of the financial crisis (see Bloom [2009]).⁵

2.4.3 Impulse Responses

Our term structure model allows us to obtain the response of the yields to shocks at all horizons, including maturities omitted in estimation. The impulse responses (*IRs*) for all maturities are known analytical functions of the parameters. Figures 2.16 and 2.17 show the *IRs*'s of 6, 12 and 60 USD- and BRL-denominated bond-yields spreads. These figures show the movements of the yield curve of different maturities (in rows) in response to one standard deviation shocks in *HY*, U.S. activity and inflation indicators and Brazilian activity, inflation, external

⁵The Economic Policy Uncertainty Index from Baker, Bloom and Davis (2012) indicate that it is very high in recent years, with clear jumps after Lehman bankruptcy, the 2010 U.S. mid-term elections, and specially the Eurozone crisis and the U.S. debt ceiling dispute.

Figure 2.17: BRL bond-yield spreads: Impulse responses from a one standard deviation shocks.



solvency and deficit indicators (in columns).

Following Bonomo and Lowenkron (2008), it is assumed that Σ is a lower triangular Cholesky matrix. The ordering of variables is the following HY , $Activity^{US}$, $Inflation^{US}$, $PC1^{US}$, $PC2^{US}$, $PC3^{US}$, $Activity^{BR}$, $Inflation^{BR}$, $Deficit^{BR}$, $ExternalSolv.^{BR}$, $PC1^{BR}$, $PC2^{BR}$, $PC3^{BR}$ - with central economy coming first. Furthermore, as a consequence of the way the principal components of yields and spreads were constructed, it is natural to impose that shocks to macroeconomic and those factors have non-correlated contemporaneous relation. Also, this kind of restriction is a standard practice in the term structure models (see Dai and Singleton [2000], Ang and Piazzesi [2003] and Joslin, Priebisch and Singleton [2010], and among others).

For the USD bond-yields spreads model, IR functions indicate that shocks to HY , U.S. activity and inflation, and Brazilian inflation and external solvency cause a statistically significant response. As expected, positive shocks to HY and

Brazilian inflation cause a positive response of spreads, while positive shocks to Brazilian external solvency cause a negative response. The short-term negative response to shocks to U.S. activity and inflation (or positive response of spreads to shocks to U.S. activity and inflation) may carry the information that negative shocks to the U.S. economy (like the one observed in 2008) increase sovereign spreads via an increase in uncertainty. We could not capture a statistically significant response of shocks to Brazilian activity and deficit.

For the BRL bond-yields spreads model, *IR* functions indicate that shocks to U.S. activity and Brazilian inflation and external solvency cause a statistically significant response. Positive shocks to Brazilian external solvency and cause a negative response, while shocks to inflation cause a positive response. Again, the short-term negative response to shocks to U.S. inflation (or positive response of spreads to shocks to U.S. inflation) may carry the information that negative shocks to the U.S. economy increase sovereign spreads via an increase in uncertainty. We could not capture a statistically significant response of shocks to *HY*, Brazilian activity and deficit.

Even though we could observe a large participation of external solvency in the reduction of both USD and BRL spreads, the relatively small response of spreads to structural shocks to this variable has a statistical explanation. Note in table 2.2 that $\Sigma_{external\ solvency^{BR}, external\ solvency^{BR}}$ is about one tenth of the analog for the other variables, indicating that the residual variance of the equation for external solvency is relatively very low. When we give a normalized one standard deviation shock in this variable, most of this shock is shrunk by the matrix Σ , and a small impact reaches spreads. It absolutely does not reduce the importance of external solvency in the composition of spreads.

2.5 Conclusion

The aim of our paper is to understand the fundamentals behind the reduction in Brazilian sovereign spreads observed in the last years in both USD and BRL bond market rather. In this sense, we are split the discount on the price of a BRL-denominated bond into a portion attributed to default risk and another portion attributed to currency risk. The first portion is estimated in the model for the Brazilian USD-denominated bonds, which is possible because of our assumption that Brazilian USD- and BRL-denominated bonds have a common default risk. The second portion is estimated in the model for the Brazilian USD-denominated bonds.

The importance of external solvency, market stress (*HY*) and local inflation (in a country suffered high inflation levels in a recent past) in the determination of recent path of spreads was expected. The evidence of Brazilian deficit (ex-interest rate payments) has a minor impact on spreads was not expected. On the one hand,

our model did not do a good job in modelling the exchange rate expectations: The coefficients on the exchange rate appreciation expectation are all statistically non-significant. This result may mean that the expected discount on the price of the defaultable BRL bond due specifically to exchange rate risk does not vary with economic fundamentals or even with yields-related variables.

On the other hand, our model did a good job in modelling USD-bond yield spreads: our measure of recovery intensity captured most of the well-known periods of stress in Brazilian bonds market. We learnt that external sector had a huge contribution for the reduction of spreads in last years and that they are not smaller because of the uncertainty environment that still lasts in the U.S. economy.

3 Geographical Heterogeneity in Regression Models

3.1 Introduction

Heterogeneity plays an important role in regression analysis. In economics, heterogeneity problems arise from two typical situations. The first one due to missing data, when the econometrician is not able to control for all relevant factors in the estimation of conditional means. If unobserved variables are correlated with the variables of interest, the inference about the regression coefficient vector is inconsistent. The second one occurs when the underlying process is not the same across the sample units. In this case, inference based on a constant coefficient vector is also inconsistent. In both cases, the availability panel data provides means of overcoming the problem, based on the potential similarity across time - either in the variable-intercept models or in the variable-coefficient model Hsiao (1989).

The goals of this paper are twofold. First, we advocate a semiparametric framework based on set of proxy variables to control for heterogeneity and unobserved effects in regression models. Contrary to Robinson (1988), who uses a kernel semiparametric correction, we consider a series (sieve) expansion of the unknown and possibly nonlinear term as in Chen (2007). Second, we show how geographical similarities can be successfully applied to control for such unobserved effects.

The use of sieve expansions has some advantages over kernel methods. First, the curse of dimensionality is not critical and multiple explanatory variables can be easily handled. Second, sieve expansions have better approximation capabilities than kernel methods, as several choices of basis functions can be shown to be dense in a given functional space. For example, in this paper we advocate the use of artificial neural network sieves which can simultaneously approximate the unknown function and its derivatives; see Hornik et al. (1994) for further details. Although, deriving the asymptotic properties for sieve extremum estimators is more complicated than the kernel case, recent results from Chen and Shen (1998), Chen and White (1998), and Chen (2007) can be easily applied in our context. Another potential difficult in semiparametric models based on sieves is the selection of the number of terms in the expansion. In this paper we show that information criteria can be used to specify the model.

We use our semiparametric model to investigate an old issue in economics -

the convergence of per capita income. The origin of the debate around convergence goes back to David Hume's 1742 essay entitled "Of the Rise and Progress of the Arts and Sciences" [Elmslie (1995)]. Whether income levels of poorer economies are growing faster than richer economies is not only an important and central question in the literature of Development Economics, but it is also related with the issue of validating competing growth theories. In the recent empirical literature, a wide array of empirical results on the subject is available, using cross-section, panel data, time-series or distribution approaches to investigate convergence both *within* an economy and *across* economies [Islam (2003)].

The paper is organized as follows. Section 3.2 describes the model and the estimation methodology. Section 3.3 presents some simulation exercises and Section 3.4 shows an application to the convergence of per capita income across Brazilian municipalities. Finally, Section 3.5 concludes. All proofs are relegated to Appendix I.

3.2 Semi-Parametric Fixed-Effects Regression Model

3.2.1 Model Definition

Consider the following assumption concerning the data generating process (DGP).

Assumption 1 (Data Generating Process). *The observed sequence of real-valued dependent variable $\{y_i\}_{i=1}^N$ is generated as*

$$y_i = a_i + \beta_0' \mathbf{x}_i + u_i, \quad i = 1, \dots, N,$$

where y_i is the dependent variable, a_i is the unobserved fixed-effect representing individual heterogeneities, and $\mathbf{x}_i \in \mathbb{R}^k$ is a set of covariates (explanatory variables) and u_i is an error term. The fixed-effects are possibly endogenous such that $\mathbb{E}[a_i | \mathbf{x}_i] \neq 0$. Furthermore, the density of \mathbf{x}_i

In this paper we propose a semiparametric approach to estimate the fixed-effects a_i , $i = 1, \dots, N$, and the parameter of interest β_0 . The core idea relies on the fact that a_i may be written as an unknown function of a set of observed and exogenous proxy variables $\mathbf{z}_i \in \mathbb{R}^q$, distinct from \mathbf{x}_i . Hence, $a_i = \eta(\mathbf{z}_i) + \epsilon_i$, where $\eta(\cdot) : \mathbb{R}^q \rightarrow \mathbb{R}$ is a unknown function. The model can be rewritten as

$$y_i = \eta(\mathbf{z}_i) + \beta_0' \mathbf{x}_i + \epsilon_i + u_i. \quad (3-1)$$

The vector $\theta = (\eta, \beta_0)'$ is defined as the parameter of interest, where η is the nonparametric nuisance parameter.¹

¹There are many ways to estimate $\eta(\cdot)$ parametrically. However, in this paper we will consider

The following assumption states two crucial conditions for the identification of (3-1).

Assumption 2 (Identification). *The following conditions hold:*

$$\mathbb{E}[y_i | \mathbf{x}_i, a_i, \mathbf{z}_i] = \mathbb{E}[y_i | \mathbf{x}_i, \mathbf{z}_i] \quad (3-2)$$

$$\mathbb{D}[a_i | \mathbf{x}_i, \mathbf{z}_i] = \mathbb{D}[a_i | \mathbf{z}_i] \quad (3-3)$$

Equation (3-2) implies that conditional on \mathbf{x}_i and \mathbf{z}_i , the expected value of y_i does not depend on a_i . This is true if \mathbf{z}_i is good enough as a proxy for a_i . If $a_i = \eta(\mathbf{z}_i) + \epsilon_i$ and $\mathbb{E}[\epsilon_i | \mathbf{x}_i] = 0$, (3-3) is trivially satisfied.

Finally, consider the following assumption about the error terms.

Assumption 3 (Errors). *Set $\xi_i = u_i + \epsilon_i$ and consider that the error sequence $\{\xi_i\}_{i=1}^N$ is formed by random variables drawn from an absolutely continuous (with respect to a Lebesgue measure on the real line), positive everywhere distribution such that $\mathbb{E}[\xi_i] = 0$, $\mathbb{E}[|\xi_i|^\delta] < \infty$, $\delta = 1, \dots, 4$, and $\mathbb{E}[\xi_i \xi_j] = 0$, $\forall i \neq j$. In addition, consider the following restrictions: (1) $\mathbb{E}[\xi_i | \mathbf{x}_i, \mathbf{z}_i] = 0$ and (2) $\mathbb{E}[\xi_i^2 | \mathbf{x}_i, \mathbf{z}_i] = \sigma_\xi^2(\mathbf{x}_i, \mathbf{z}_i) < \infty$, $\forall i$.*

3.2.2 Estimation Method and Asymptotic Theory

The key idea of this paper is to estimate jointly both the parametric and the nonparametric components of $\boldsymbol{\theta}$ by the sieve extremum estimation method.

Set $\rho(\mathbf{v}_i; \boldsymbol{\theta}) \equiv \rho(\mathbf{v}_i; \boldsymbol{\beta}, \eta(\cdot)) = y_i - \boldsymbol{\beta}' \mathbf{x}_i - \eta(\mathbf{z}_i)$, where $\mathbf{v}_i = (y_i, \mathbf{z}_i', \mathbf{x}_i')'$, $i = 1, \dots, N$. Assume that the conditional expectation $\mathbb{E}[\rho(\mathbf{v}_i; \boldsymbol{\theta}_0) | \mathbf{x}_i, \mathbf{z}_i] = 0$, where $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\beta}_0', \eta_0)' \in \Theta$ is the true parameter vector. Furthermore, define $\sigma^2(\mathbf{v}_i) = \mathbb{E}[\rho(\mathbf{v}_i; \boldsymbol{\theta})^2 | \mathbf{x}_i, \mathbf{z}_i]$. Let $\Theta = \mathcal{B} \times \mathcal{H}$, where \mathcal{H} is a space of continuous functions defined on a bounded set of \mathbb{R}^q and \mathcal{B} is a compact set in \mathbb{R}^k . Consider also a sequence of approximating parameter spaces (or sieves) as $\Theta_N = \mathcal{B} \times \mathcal{H}_N$, where $\bigcup_N \mathcal{H}_N$ is dense in \mathcal{H} in some desirable metric.

To obtain an efficient estimator of $\boldsymbol{\beta}_0$ we apply the following three-step procedure as suggested by Ai and Chen (2003):

Step 1 Obtain an initial consistent sieve nonlinear least squares estimator $\hat{\boldsymbol{\theta}}_N = (\hat{\boldsymbol{\beta}}_N, \hat{\eta}_N)$ by

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{(\boldsymbol{\beta}, \eta) \in \mathcal{B} \times \mathcal{H}_N} \frac{1}{N} \sum_{i=1}^N \rho(\mathbf{v}_i; \boldsymbol{\theta})^2.$$

Step 2 Obtain a consistent estimator $\hat{\sigma}^2(\mathbf{v}_i)$ of $\sigma_0^2(\mathbf{v}_i) \equiv \mathbb{E}[\rho(\mathbf{v}_i; \boldsymbol{\theta}_0)^2 | \mathbf{x}_i, \mathbf{z}_i]$ using $\hat{\boldsymbol{\theta}}_N = (\hat{\boldsymbol{\beta}}_N, \hat{\eta}_N)$.

solely the nonparametric alternative.

Step 3 Obtain the optimally weighted estimator $\tilde{\boldsymbol{\theta}}_N = (\tilde{\boldsymbol{\beta}}_N, \tilde{\eta}_N)$ by solving

$$\tilde{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta} \in \mathcal{B} \times \mathcal{H}_N} \mathcal{Q}(\boldsymbol{\theta}, M),$$

$$\text{where } \mathcal{Q}(\boldsymbol{\theta}, M) \equiv \frac{1}{N} \sum_{i=1}^N \frac{\rho(\mathbf{v}_i; \boldsymbol{\theta})^2}{\hat{\sigma}^2(\mathbf{v}_i)}.$$

There are a number of distinct sieve estimators. In this paper we advocate the use of the Artificial Neural Network (ANN) sieve defined as

$$\eta_0(\mathbf{z}_i) \in \mathcal{H} \equiv \left\{ \alpha_0 + \sum_{m=1}^{M_N} \alpha_m f(\mathbf{z}_i; \boldsymbol{\omega}_m, c_m) \right\},$$

where

$$f(\mathbf{z}_i; \boldsymbol{\omega}_m, c_m) = \frac{1}{1 + e^{-(\boldsymbol{\omega}'_m \mathbf{z}_i - c_m)}},$$

and $|\alpha_0| < \infty$, $|\alpha_m| < \infty$, $|c_m| < \infty$, and $\|\boldsymbol{\omega}_m\| < \infty$, $m = 1, \dots, M_N$. The class of ANN sieves is dense in \mathcal{H} .

Consider the following assumption.

Assumption 4 (Approximation Capability). *There exists a small number $\delta > 0$ such that $|a_i - \eta_0(\mathbf{z}_i)| < \delta$, $\forall i$, a.s..*

Assumption 4 states that the unobserved fixed-effects may be approximated arbitrarily accurately by a function of \mathbf{z}_i , $i = 1, \dots, N$. As the class of ANN sieves is dense in \mathcal{H}_N , and thus is an universal approximator to the unknown function $\eta_0(\mathbf{z}_i)$, the semi-parametric estimator $\hat{\eta}$ can be used to control for unobserved characteristics, leading to an unbiased estimator of $\boldsymbol{\beta}$.

Define $\mathbf{D}(\mathbf{x}_i, \mathbf{z}_i) = \mathbf{x}_i - \mathbb{E}[\mathbf{x}_i | \mathbf{z}_i]$ and $\mathbf{D}(\mathbf{X}, \mathbf{Z}) = [\mathbf{D}(\mathbf{x}_1, \mathbf{z}_1) \dots \mathbf{D}(\mathbf{x}_N, \mathbf{z}_N)]'$. Following Ai and Chen (2003), the next proposition states the asymptotic distribution of $\tilde{\boldsymbol{\beta}}$.

From the results in Ai and Chen (2003) and under Assumptions 1–4, it follows that

$$\sqrt{N}(\tilde{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0) \xrightarrow{d} \mathbf{N}(0, \mathbf{G}_0^{-1}),$$

with

$$\mathbf{G}_0 = \mathbb{E} \left\{ \mathbf{D}(\mathbf{X}, \mathbf{Z})' \left[\hat{\boldsymbol{\Sigma}}(\mathbf{X}, \mathbf{Z}) \right]^{-1} \mathbf{D}(\mathbf{X}, \mathbf{Z}) \right\}$$

and $\boldsymbol{\Sigma}(\mathbf{X}, \mathbf{Z}) = \text{diag} [\sigma_0^2(\mathbf{v}_1), \dots, \sigma_0^2(\mathbf{v}_N)]$.

3.2.3 Model Selection

In applications the number of sieves is unknown and should be determined from the data. In the neural network literature several approaches have been proposed. A popular method for doing that is pruning, in which a model with a large

number of hidden units is estimated first, and the size of the model is subsequently reduced by applying an appropriate technique such as cross-validation. Another technique used in this connection is regularization, which may be characterized as penalized maximum likelihood or least squares applied to the estimation of neural network models. Bayesian regularization, based on selecting a prior distribution for the parameters, may serve as an example.

Another possibility, which is adopted in the paper, is to estimate R models, with $M_N = 1, \dots, R$, for R sufficiently large, and choose the optimal M_N^* based on the use of model selection criteria (MSC). In this section we show the consistency of an IC to determine the number of nonlinear terms in our model (M_N). Collect the data points in a vector \mathbf{V} . Define $M \in \{0, 2, \dots, \bar{M}\}$ and consider a class of models $\mathcal{M}(\mathbf{V}; \boldsymbol{\theta}_M)$ indexed by M . The parameter vector $\boldsymbol{\theta}_M$ is indexed by M . Denote M^* as the true value of M . Our goal is to estimate M by minimizing the following IC:

$$\text{IC}(M) = \mathcal{Q}(\boldsymbol{\theta}, M) + \lambda_N(M), \quad (3-4)$$

where $\mathcal{Q}(\boldsymbol{\theta}, M)$ is a cost function as defined before and $\lambda_N(M)$ is a penalty term to be defined later.

Define the sets

$$\Theta_M^* = \{\boldsymbol{\theta} \in \Theta : \boldsymbol{\theta} = \arg \min \mathbb{E}[\mathcal{Q}(\boldsymbol{\theta}, M)]\} \text{ and } \Theta_{T,M} = \{\boldsymbol{\theta} \in \Theta : \boldsymbol{\theta} = \arg \min \mathcal{Q}(\boldsymbol{\theta}, M)\}.$$

Assumption 5 (Class of Models and Penalty Function). *Assume that:*

1. $\mathcal{M}(\mathbf{V}; \boldsymbol{\theta}_1) \subseteq \mathcal{M}(\mathbf{V}; \boldsymbol{\theta}_2) \subseteq \dots \subseteq \mathcal{M}(\mathbf{V}; \boldsymbol{\theta}_{\bar{M}})$;
2. $\lambda_N(M)$ is a positive and increasing function of M ;
3. $\frac{1}{N}\lambda_N(M) \rightarrow 0$ and $\lambda_N(M) \rightarrow \infty$, as $N \rightarrow \infty$ for every $M \in \{0, 2, \dots, \bar{M}\}$;
4. \bar{M} is such that

$$\sum_{M=M^*+1}^{\bar{M}} \left[\frac{k_{\theta,M}}{\lambda_N(M) - \lambda_N(M^*)} \right] \rightarrow 0, \text{ as } N \rightarrow \infty,$$

where $k_{\theta,M}$ is the dimension of $\boldsymbol{\theta}$, i.e., $k_{\theta,M} = k + q$;

5. for every $M = 0, 1, 2, \dots, \bar{M}$ and some positive constant c

$$\max_{\hat{\boldsymbol{\theta}} \in \Theta_{N,M}} \min_{\boldsymbol{\theta} \in \Theta_M^*} \mathbb{E} \left| \mathcal{Q}(\hat{\boldsymbol{\theta}}, M) - \mathcal{Q}(\boldsymbol{\theta}, M) \right| \leq ck_{\theta,M}^2 \text{ with probability 1 as } N \rightarrow \infty.$$

Assumptions 5(1)-(4) define how the penalty term can be specified as well as the order of increase in the number of candidate models. Assumption 5(5) requires that the elements in $\Theta_{N,M}$ and Θ_M^* are arbitrarily close.

Theorem 1. Under Assumptions 1–5, $P(\widehat{M} \neq M^*) \rightarrow 0$, as $N \rightarrow \infty$.

3.3 Simulations

3.3.1 Setup

In this section we conduct a Monte Carlo study to check the finite sample properties of the estimator discussed in this paper and compare it with alternatives available in the literature. We simulate the following DGP:

$$y_i = a_i + x_i + u_i,$$

where $x_i \in \mathbb{R}$ and

$$a_i = f(\mathbf{z}_i; \zeta) = \sum_{j=1}^{\zeta} \left(\sum_{k=1}^K z_{ik} \right)^j + \varepsilon_i,$$

and \mathbf{z}_i is a K -dimensional vector. The parameter ζ is a complexity index. $\{x_i, \mathbf{z}_i\}_{i=1}^N$ is generated from a normal distribution, such that:

$$(x_i, \mathbf{z}'_i)' \equiv (x_i, z_{1i}, z_{2i}, z_{3i})' \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad i = 1, \dots, N,$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 2.5 & \cdot & \cdot & \cdot \\ -0.3 & 1 & \cdot & \cdot \\ 1 & -0.2 & 1.6 & \cdot \\ 1 & -0.3 & -0.1 & 1.3 \end{pmatrix}.$$

The disturbances u_i and ε_i are generated from independent standard normal distributions. Simulations with 400 repetitions are performed using different combinations of values for ζ , K and N : $\zeta \in \{2, 3, 4\}$, $K \in \{1, 2, 3\}$, $N \in \{200, 1000, 3000\}$. The sieve estimation procedure was applied to the data with identity weighting matrix $\widehat{\boldsymbol{\Sigma}}_0(\mathbf{X}, \mathbf{Z}) = \mathbf{I}$. Three competing procedures are also evaluated: (1) ordinary Least Squares Model omitting a_i (OLS-OV), $y_i = x_i\beta + \xi_i$; an Ordinary Least Squares Model including as regressors (x_i, \mathbf{z}'_i) (OLS), $y_i = \beta x_i + \boldsymbol{\gamma}'\mathbf{z}_i + \xi_i$; and finally a version of Robinson (1988) estimation method.

Shortly, our version of Robinson (1988) estimation procedure consists of the following steps: (1) Regress x on \mathbf{z} using a nonparametric estimation method and collect the vector of residuals \mathbf{U}_X ; (2) Regress \mathbf{Y} on \mathbf{Z} using a nonparametric estimation method and collect the vector of residuals \mathbf{U}_Y ; (3) Obtain $\widetilde{\boldsymbol{\beta}} = (\mathbf{U}'_X \mathbf{U}_X)^{-1} \mathbf{U}'_X \mathbf{U}_Y$. In steps (1) and (2), residual estimation are based on a Nadaraya-Watson kernel estimator with the Gaussian kernel.

In this exercises, the number of sieves M_N is chosen through the minimization of the Hannan-Quinn information criterion (HQC) in a range between 0 (linear case) and 15 sieves.

3.3.2 Results

The median, mean and standard error (SE) of each of the estimators of β across the 400 simulations are reported in tables 3.1-3.3. In each table, we present the results for the competing estimation methods: OLS-OV, OLS, Robinson and sieves for different values for the complexity index ζ and the number of observations N .

When the number K of variables in the semi-parametric component is one (table 3.1), Robinson's (1988) estimation method has a good performance. When the complexity index ζ equals 2, the mean of estimated parameter ranges between 0.98 ($N = 200$) and 1.00 ($N = 3000$). When the complexity increases to four, Robinson's estimator has a worse performance, but is still not too far from the true value of β , especially when the sample is large: for $N = 3000$ and $\zeta=4$, the mean of Robinson's estimator is 0.98. The means of our sieve point estimations are all either 1.00 or 1.01 and the median of the number of sieves is between 2 and 4. OLS and OLS-OV estimators have a poor performance in all exercises with different combinations of number of observations and complexity indices.

Table 3.2 shows the simulation results for the two variables case ($K = 2$). In this case, the performance of Robinson's estimator gets dramatically worse as the complexity index increases. As already noted in Robinson (1988), the performance of this estimator is poor when the number of variables in the semi-parametric component is larger than one. On the other hand, our sieve estimator keeps on having a good performance: in all exercises with different combinations of number of observations and complexity indices, its the mean and median is between 1.00 and 1.01. The median of the number of sieves increased for a range between 3 and 8. OLS and OLS-OV estimators have an even poorer performance in the 2 variable case than in the 1 variable case.

The three variable case results are displayed in table 3.3 and have a similar general picture. The sieve estimator is really close to one, which is the real value of β , and Robinson, OLS and OLS-OV estimators have poor performances, getting worse as the complexity index increases.

3.4 Applications: Economic Growth and Convergence

This section illustrates the use of our semi-parametric regression model to test convergence. We use latitude and longitude data attempting to control for unobserved characteristics among Brazilian municipalities. Our starting point is a

textbook Solow model featuring the Cobb-Douglas production function with labor-augmenting technological progress:

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha},$$

where Y_t is output, K_t is capital, and $A_t L_t$ is effective unit of worker. L and A are assumed to grow exogenously at rates n and g so that

$$L_t = L_0 e^{nt}, \quad \text{and} \quad A_t = A_0 e^{gt}.$$

Let \hat{y}^* be the steady state level of income per effective worker, and let \hat{y}_t be its actual value at any time. The convergence equation is given by (Islam [1995]):

$$\dot{\ln \hat{y}_t} = \lambda (\ln \hat{y}^* - \ln \hat{y}_t),$$

where $\lambda = (1 - \alpha)(n + g + \delta)$, and δ is the depreciation rate. This equation implies that:

$$\ln \hat{y}_{t_2} - \ln \hat{y}_{t_1} = (1 - e^{-\lambda\tau})(\ln \hat{y}^* - \ln \hat{y}_{t_1})$$

where $\tau = t_2 - t_1$. Substituting $\ln \hat{y}_t = \ln y_t - \ln A_0 - gt$ and the parameters that determine the steady state output per capita y^* gives (Islam [1995]):

$$\begin{aligned} \ln y_{t_2} - \ln y_{t_1} = & (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln s \\ & - (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln(n + g + \delta) \\ & - (1 - e^{-\lambda\tau}) \ln y_{t_1} \\ & + (1 - e^{-\lambda\tau}) \ln A_0 + g(t_2 - e^{-\lambda\tau} t_1), \end{aligned}$$

where s is the saving rate and A_0 is the initial level of the labor-augmenting technology.

3.4.1 Brazilian Municipalities

We illustrate the use of our semi-parametric regression model to test convergence among Brazilian municipalities between 1970 and 2000 controlling for unobserved characteristics. Our starting point is the simplified version of equation (4-5) presented in Barro and Sala-i-Martin (1992):

$$\log \left(\frac{y_{i,t}}{y_{i,t-1}} \right) = a_i + \gamma \cdot \log y_{i,t-1} + \phi_i \cdot (t - 1) + u_{i,t}, \quad (3-5)$$

where $y_{i,t}$ is the per capita income of region i in period t , a_i is associated with the steady-state level of per capita income and the rate of technological progress, ϕ_i is

a parameter related to the time trend determined by the technological progress, and $u_{i,t}$ is the random term. Convergence corresponds to the parameter γ .

From a conceptual point of view, there are two alternative assumptions determining the most important distinction of convergence concepts. First, we can assume that $a_i = a$ and $\phi_i = \phi$, i.e., the basic parameters of preference and technology are the same for all economies represented in the sample. This is the case where $\gamma < 0$ represents *unconditional convergence* - a situation where poor economies tend to grow unconditionally faster than rich ones. Alternatively, we can state a weaker assumption allowing for possible differences in the steady state across the economies considered. In terms of equation (3-5), a_i and ϕ_i are allowed to vary. In this case, $\gamma < 0$ means *conditional convergence* - controlling for differences in the steady-state per capita income, poor economies grow faster.

Here, we estimate (3-5) in a cross-section setup, where there is no identifiable time trend and we are not able to distinguish between ϕ_i and a_i . Thus, we estimate the following equation:

$$\log \left(\frac{y_{i,2000}}{y_{i,1970}} \right) = \alpha_i + \gamma \cdot \log y_{i,1970} + u_{i,2000}. \quad (3-6)$$

Our data comes from the Brazilian Demographic Censuses of 1970 and 2000. The geographical units were adapted to incorporate the changes in the organization of the Brazilian territory in the period. In 1970, Brazil comprised 3,951 municipalities which, in 2000, became 5,507. Therefore, all the information collected in 2000 were aggregated in order to match the municipal structure of 1970. Our dependent variable is the average per capita income growth between 1970 and 2000 for each city. The independent variable is the logarithm of the per capital income level in 1970. These two variables are presented in Figures 1 and 2.

The first characteristic depicted in Figures 1 and 2 is the large differences across municipalities, both in terms of growth rate or 1970 income level. Figure 2 shows a sharp contrast between the poor Northeastern part of the country and the Southeast and South regions. The second characteristic is that the variation in both growth and income levels do not coincide with the administrative State frontiers. There is a lot of variation within many of the Brazilian States.

We consider different formulations in the estimation of equation (3-6). For the case of unconditional convergence, where $\alpha_i = \alpha$, equation (3-6) can be estimated by OLS, providing consistent estimates for γ if $E(\log y_{i,1970} \cdot u_{i,2000}) = 0$. This result is reported in column (1) of table 3.4 and in the scatter plot of Figure 3. The estimated γ coefficient is -0.004, significant at the significance level of 1%. Thus, the poorer municipalities in 1970 experienced a (unconditionally) lower growth rate between 1970 and 2000.

Although we found significant evidence of unconditional convergence in

Brazilian municipalities, Brazil is a large country that displays huge regional differences as shown in Figures 1 and 2. As a consequence, we might expect significant variation in the steady-state levels of per capita income across cities. Therefore, the concept of conditional convergence seems more plausible for our sample.

However, the study of conditional convergence in this cross-section environment, is a more complex task. On the one hand, we use no additional data to approximate α_i . On the other hand, there is no degree of freedom to estimate α_i without additional statistical structure.

A natural strategy is to use aggregation through dummy variables to enable the estimation of different steady-states, based on administrative divisions such as the 5 Brazilian macro-regions or the 27 States. The drawback of this approach is that administrative divisions of the territory do not necessarily represent the differences in the steady-state levels of per capita income and technology. Indeed, Figures 1 and 2 show that many States have important *within* variation.

Columns (2) and (3) report the estimation of equation (3-6), controlling for dummy variables regarding the Brazilian macro-regions and states, respectively. In both cases, the estimated coefficient of convergence indicates a stronger process of conditional convergence, where $\hat{\gamma} = -0.014$ is significant at 1%.

Next, we use the semi-parametric approach presented in the previous section to evaluate conditional convergence. Differences in the preference and technological parameters are endogenously incorporated into analysis through geographical similarities. The underlying assumption is that nearby cities face similar steady-states. In terms of the modeling, we estimate (3-6), considering that α_i is a semi-parametric function of the latitude and longitude coordinates, as in (3-1). The result is presented in the column (4) of table 3.4, while the estimated α_i are depicted in Figure 3.5.

The semi-parametric approach give us a highly significant $\hat{\gamma} = -0.017$, which is also statistically different from the OLS point estimate -0.014 . The flexibility of having endogenously determined geographical differences across steady-state levels illustrates that bias produced by the OLS approach based on inadequate state dummies. But, although the difference is statistically important, it is not economically significant.

The most relevant effect of the semi-parametric approach to the study of convergence in our sample can be viewed in Figure 3.4, which present the plot of the growth rate adjusted by differences in the steady-state $\left(\log \left(\frac{y_{i,2000}}{y_{i,1970}} \right) - \alpha_i \right)$ against the logarithm of 1970 per capita income. Comparing with Figure 2, we observe a clearer linear pattern. Controlling for differences in the steady-state, poor cities growth faster than rich cities.

Table 3.1: Simulation Results: $K=1$ variable and different combinations of number of observations N and complexity ζ .

		N = 200			N = 1000			N = 3000		
		$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$
Method	Moment	β estimates								
OLS-OV	Median	0.647	(0.020)	(1.817)	0.646	(0.061)	(1.897)	0.638	(0.089)	(2.014)
	Mean	0.644	(0.056)	(1.929)	0.645	(0.069)	(1.935)	0.638	(0.088)	(2.021)
	SE	0.214	0.595	1.872	0.096	0.269	0.849	0.055	0.156	0.502
OLS	Median	0.997	0.896	0.618	1.003	0.886	0.580	1.002	0.868	0.497
	Mean	0.999	0.888	0.543	1.003	0.882	0.557	1.000	0.875	0.507
	SE	0.091	0.281	1.145	0.040	0.132	0.546	0.025	0.082	0.340
Robinson	Median	0.983	0.929	0.746	0.998	0.966	0.882	0.998	0.983	0.935
	Mean	0.982	0.926	0.739	0.996	0.966	0.884	0.998	0.982	0.935
	SE	0.067	0.109	0.415	0.028	0.038	0.112	0.016	0.018	0.042
Sieves	Median	0.996	0.998	1.006	1.001	0.996	0.998	1.000	0.999	1.000
	Mean	0.996	1.000	1.004	1.002	0.996	0.999	1.001	0.998	1.000
	SE	0.066	0.067	0.065	0.028	0.032	0.028	0.016	0.016	0.018
Method	Moment	M_N estimates (Hannan-Quinn Information Criteria choice in a range between 0 and 15)								
Sieves	Median	2.000	3.000	3.000	2.000	3.000	3.000	2.000	3.000	4.000
	Mean	2.560	3.310	3.308	2.503	3.463	3.525	2.488	3.438	3.770
	SE	1.277	1.154	1.091	0.918	0.878	0.788	0.819	0.814	0.961

Figure 3.5 shows that there are significant geographical differences across municipalities which do not coincide with the geographical structure of the Brazilian States, especially in those of the Southeast and South regions. This is exactly what explain the statistical difference between the OLS estimates and the parameter obtained in our framework.

3.5 Conclusion

This paper has proposed a semi-parametric approach to control for unobserved fixed-effects in linear regression models. A methodology based on the Artificial Neural Network sieve extremum estimator has been advocated. The semi-parametric fixed-effect regression model is applied to the study of convergence across Brazilian municipalities for the period from 1970 to 2000. The estimated fixed-effects account for differences in the steady-state levels of per capita income, allowing for a more evident convergence relation. The estimated coefficient of convergence is statistically different from the OLS counterparts, and the relationship between the (adjusted) growth rate and the per capita income level becomes closer to a straight and negatively sloped line. We estimate strong evidence of conditional convergence among Brazilian cities between 1970 and 2000.

Table 3.2: Simulation Results: $K=2$ variables and different combinations of number of observations N and complexity ζ .

		N = 200			N = 1000			N = 3000		
		$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$
Method	Moment	β estimates								
OLS-OV	Median	4.205	28.093	196.863	4.422	29.335	210.768	4.381	29.371	212.203
	Mean	4.274	28.516	205.336	4.365	29.268	210.864	4.378	29.387	211.862
	SE	1.222	9.675	79.199	0.597	4.673	36.707	0.353	2.783	21.979
OLS	Median	2.739	21.234	167.701	2.736	21.694	177.650	2.728	21.657	177.608
	Mean	2.738	21.671	178.029	2.733	21.578	177.167	2.734	21.636	177.780
	SE	0.402	5.514	60.798	0.182	2.373	24.862	0.102	1.344	14.383
Robinson	Median	1.608	6.829	43.410	1.337	4.854	31.410	1.207	3.659	22.698
	Mean	1.621	7.022	46.146	1.337	4.849	31.242	1.209	3.687	22.907
	SE	0.276	3.190	28.910	0.078	0.946	8.953	0.035	0.446	4.412
Sieves	Median	1.004	1.001	1.006	0.999	1.003	1.000	1.001	1.000	1.001
	Mean	1.004	1.002	1.004	0.998	1.003	1.001	1.001	1.001	1.002
	SE	0.081	0.080	0.082	0.035	0.034	0.032	0.019	0.020	0.019
Method	Moment	M_N estimates (Hannan-Quinn Information Criteria choice in a range between 0 and 15)								
Sieves	Median	4.000	5.000	7.000	3.000	4.500	8.000	4.000	4.000	7.000
	Mean	5.165	5.903	7.723	3.778	5.290	8.438	4.095	4.680	7.250
	SE	3.535	3.530	3.699	2.242	2.574	2.775	2.083	2.461	2.291

Table 3.3: Simulation Results: $K=3$ variables and different combinations of number of observations N and complexity ζ .

		N = 200			N = 1000			N = 3000		
		$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$	$\zeta = 2$	$\zeta = 3$	$\zeta = 4$
Method	Moment	β estimates								
OLS-OV	Median	17.622	238.881	2856.905	17.121	236.053	2856.514	17.267	237.283	2877.016
	Mean	17.346	237.417	2867.028	17.212	236.390	2861.914	17.307	237.788	2880.014
	SE	4.110	49.613	588.705	1.728	21.167	254.812	1.070	13.092	156.859
OLS	Median	9.402	164.712	2239.662	9.448	164.943	2248.778	9.463	165.549	2259.428
	Mean	9.433	164.799	2245.658	9.459	165.431	2257.087	9.476	165.921	2265.451
	SE	0.764	18.785	324.801	0.369	9.038	154.257	0.209	5.107	88.062
Robinson	Median	9.923	137.307	1636.507	7.242	101.535	1217.770	5.695	80.278	966.420
	Mean	10.092	139.702	1660.477	7.261	101.738	1218.606	5.693	80.354	967.781
	SE	1.362	22.220	293.609	0.453	7.827	104.844	0.219	3.903	51.905
Sieves	Median	1.011	1.000	1.003	1.006	0.997	0.998	0.997	0.998	0.998
	Mean	1.008	0.996	1.006	1.004	0.998	0.999	0.998	0.998	0.999
	SE	0.111	0.115	0.119	0.046	0.047	0.048	0.027	0.026	0.024
Method	Moment	M_N estimates (Hannan-Quinn Information Criteria choice in a range between 0 and 15)								
Sieves	Median	5.000	5.000	7.500	4.000	3.000	6.000	4.000	3.000	6.000
	Mean	7.113	6.960	8.180	5.275	4.500	6.745	5.120	4.465	6.738
	SE	4.829	4.624	3.924	3.786	3.241	3.137	3.184	3.092	2.787

Table 3.4: Convergence Regressions - Brazilian Municipalities.
Dependent variable: per capita growth rate, 1970-2000

	(1)	(2)	(3)	(4)
log(per capita income level 1970)	-0.004*** (0.000)	-0.014*** (0.000)	-0.014*** (0.000)	-0.017*** (0.000)
Constant	0.059*** (0.001)	0.104*** (0.002)	0.118*** (0.009)	- -
Number of Sieves	-	-	-	12
Macrorregion dummies (5 regions)	No	Yes	No	No
State dummies (27 states)	No	No	Yes	No
Method of estimation	OLS	OLS	OLS	Sieves GLS
Observations	3948	3948	3948	3948
R-squared	0.031	0.386	0.487	0.498

Note: Standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 3.1: Map of the log(Brazilian per capita income growth 1970-2000)

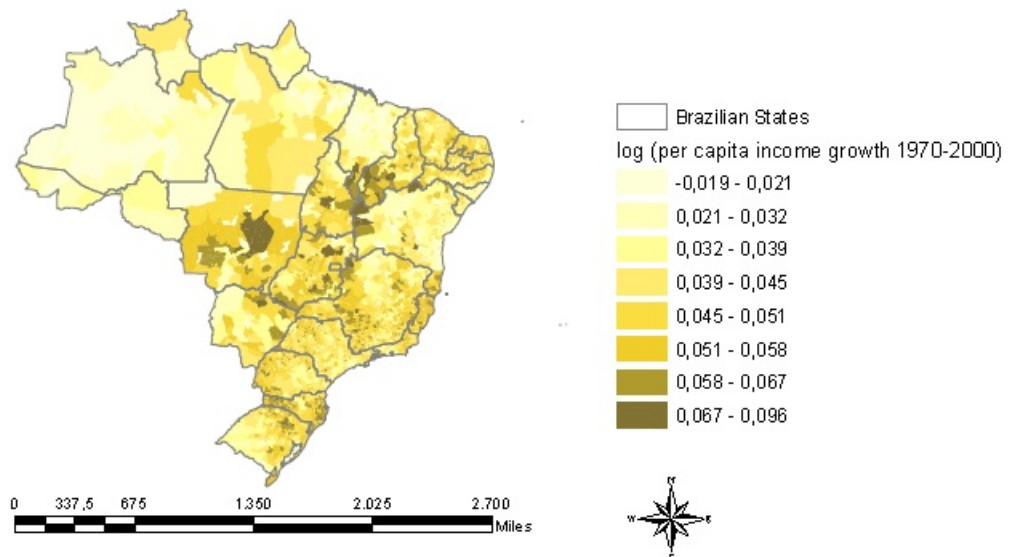


Figure 3.2: Map of the log(Brazilian per capita income level 1970)

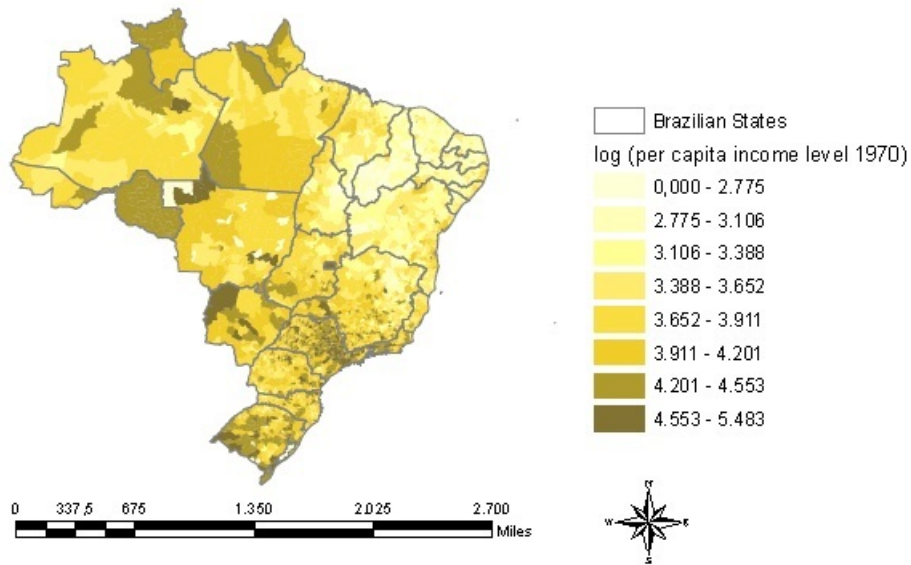


Figure 3.3: Growth rate from 1970 to 2000 vs. log(1970 per capita income level)

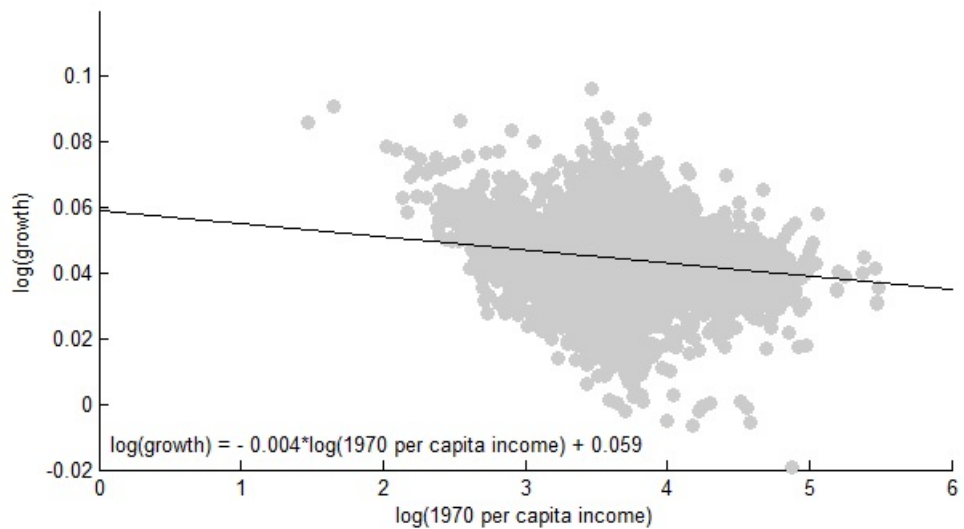


Figure 3.4: Conditional Growth rate from 1970 to 2000 vs. log(1970 per capita income level)

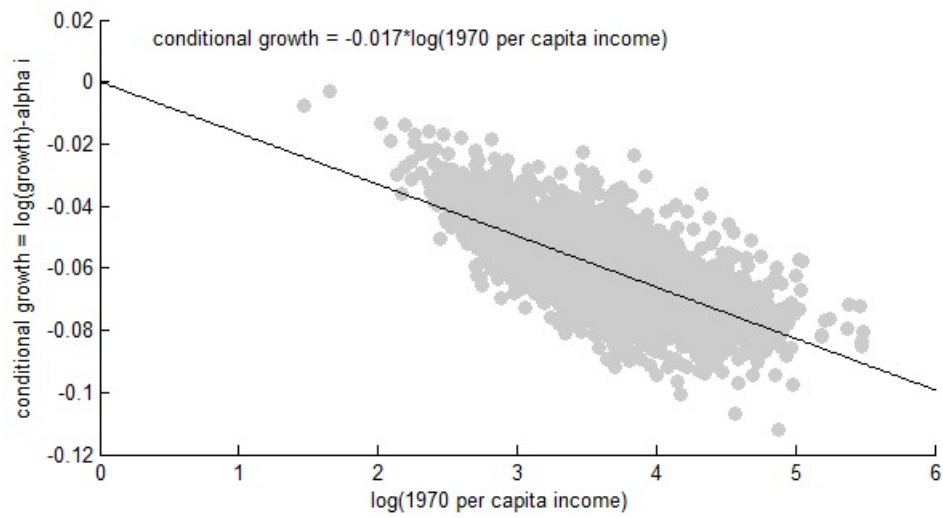
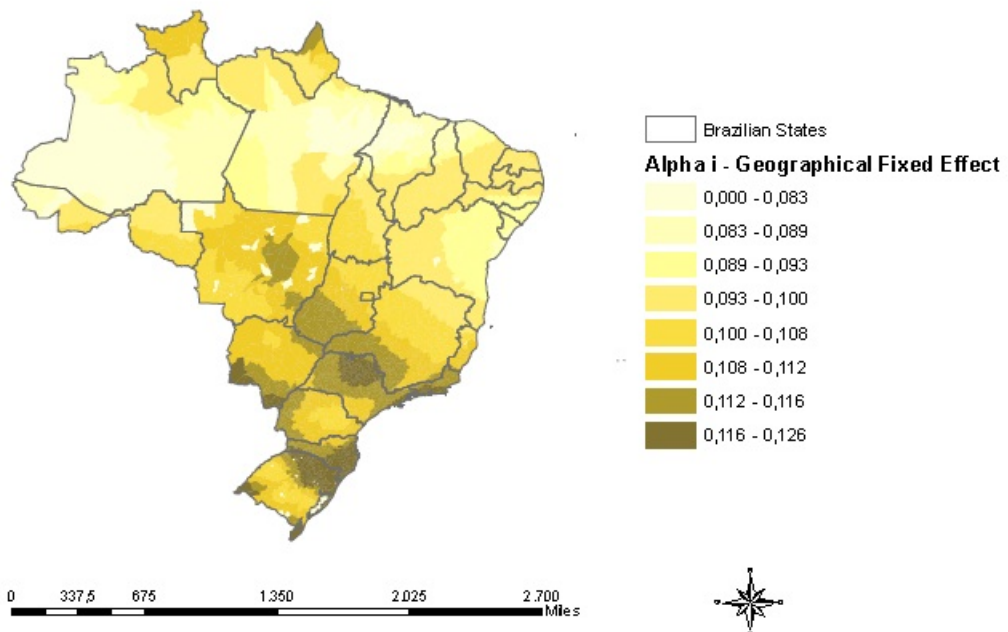


Figure 3.5: Map of the Semi-Parametric Geographical Fixed Effects



4 Is the Convergence of the Manufacturing Sector Unconditional?

4.1 Introduction

The rapid pace of economic growth that emerging and developing economies experienced in the last decades, specially in the run-up to the global financial crisis of 2008-2009, has given a new life to the debate about economic convergence - i.e., whether poorer countries tend to grow faster than the richer ones, then converging in living standards. Discussions about the risk of decay of the supremacy of the U.S. and other advanced economies (Eichengreen 2011, Subramanian 2011), and the prospects of the developing world growth (O'Neill 2011, Rodrik 2011a) abound.

Rodrik (2011a, b) documented that manufacturing industries exhibit unconditional convergence in labor productivity. Using the same data set - UNIDO's INDSTAT 4, available for a wide range of countries -, we provide a novel semi-parametric specification for convergence equations and show that the pace of convergence varies systematically with country-specific characteristics. The semi-parametric approach we propose identify unobserved heterogeneity in the convergence coefficient through geographic, political and educational indicators. We will consider the flexible smooth transition model with multiple groups and multiple transition variables proposed by Medeiros and Veiga (2005). This model allows that each group has distinct dynamics controlled by a linear combination of known variables. This formulation can be interpreted as coefficient varying linear model where the coefficients are the outputs of a hidden layer feedforward neural network.

We found evidence that the laws of motion for growth are different across countries and those with worse institutions converge faster. The convergence coefficient also has a non-monotonic relationship with trade openness and education, being faster at the extremes. The differences in the convergence coefficient across countries is not only of statistical significance, but it is also economically meaningful. The extreme values of the estimated convergence coefficient in are -3.7% and -2.8% per year, which means that the half life to productivity convergence varies in a range of about 10 years (between 27 and 37 years).

4.1.1 Literature

Whether income levels of poorer economies are growing more rapidly than richer economies is not only an important question in the literature of Development Economics, but it is also related with the issue of validating competing growth theories. In the neoclassical growth literature, unconditional convergence implies that there is only one steady state level of per capita income to which all economies approach, and conditional convergence implies that equilibrium differs by economy, and each particular economy approaches its own but unique per capita income equilibrium (Islam 2003). There are numbers of works with different approaches showing evidences of conditional convergence (Mankiw, Romer, and Weil 1992 and Islam 1995). It is widely known, however, that empirical works have found hard to prove unconditional economic convergence when a broad and diversified sample of countries is considered (Islam 2003, Durlauf, Johnson and Temple 2005).

Baumol (1986) shows that (unconditional) convergence of output per capita is observed among developed countries, but it is not shared by less developed economies, suggesting that there would exist “convergence clubs”. Indeed, a non-linear specification for the growth equation hold for a class of growth models, starting with Azariadis and Drazen (1990). Their model produces multiple locally stable steady states in per capita output. Cross-country growth behavior in these models exhibits multiple regimes as countries associated with the same steady state obey a common linear regression.

Durlauf and Johnson (1995) and Sachs and Warner (1995) explore it dividing a sample of develop and developing countries in groups based on country characteristics . They show that the laws of motion for growth within each subgroup are different: in growth regressions, the estimated coefficient on the initial level of GDP per capita (the convergence coefficient), although always negative, changes substantially, and is not statistically significant in all cases.

More recently, Canova (2004) proposed a Bayesian procedure to examine the likelihood of convergence clubs in the distribution of income per capita. The break points are identified through the ordering of observations according to country characteristics, and this method allows him to identify clubs and estimate the convergence coefficient of each club. But, we still cannot access *how* each of these variables are related to the converge coefficient.

In recent works, Rodrik (2011a, b) gave a new breath to the debate about convergence. His works suggest that we can find unconditional convergence if we look at industries instead of the whole economy. He documents evidence of unconditional convergence in 4-digit manufacturing industries for a large group of develop and developing countries over a period since 1990. Since unconditional convergence implies the existence of only one income per capita equilibrium level

to be shared among all economies, is quite intuitive that it is true in sectors that face more external influence.

Hwang (2007, chap. 3) has documented that there is a tendency for unconditional convergence in export unit values in highly disaggregated product lines. Hwang shows that the lower the average unit values of a country's manufactured exports, the faster the country's subsequent growth, unconditionally. If there is unconditional convergence in unit values of exports, it may be true that the convergence coefficient may vary across countries, depending on, for instance, the openness of the economy.

Rodrik (2011a, b) presented evidence that the productivity growth of low-productivity industries is larger. He also suggests that the industry-level unconditional convergence is not uniform across manufacturing industries, i.e., the pace of convergence changes across industries. There would exist a hierarchy within manufacturing - the convergence would be most rapid in machinery and equipment and least rapid in textiles and clothing. In this paper, we bring evidence that pace of convergence changes across countries. Our basic questions are: Can we identify a multiple regime dynamics in industry productivity growth across countries? Do country-specific features are related to the industry productivity growth? In what magnitude? Is the signal the expected one? This is what this paper is about.

Many empirical attempts to identify different dynamics of growth across countries have been made (Baumol 1986, Durlauf and Johnson 1995, Sachs and Warner 1995, Canova 2004). None of them allow us to access how variables used to group countries with common growth dynamics are related to the growth dynamics itself. It is important, for example, to access to what extent countries that adopt sound policies have been awarded with higher growth.

Instead of splitting samples (as in Baumol 1986, Durlauf and Johnson 1995 and Sachs and Warner 1995), we will allow the convergence coefficient to vary across countries, and this variation will depend on geographic, political and educational indicators. We propose a semi-parametric approach to identify unobserved heterogeneity in the convergence coefficient through these indicators. We will consider the flexible smooth transition model with multiple groups (or time regimes) and multiple transition variables proposed by Medeiros and Veiga (2005). This model allows that each group has distinct dynamics controlled by a linear combination of known variables such as geographic, political and educational indicators. This formulation can be interpreted as coefficient varying linear model where the coefficients are the outputs of a hidden layer feedforward neural network. The model is estimated using a sieve extremum estimator which is shown to be dense in a given functional space. We use an Artificial Neural Network sieve, which has optimal convergence properties.

There are at least two advantages of this approach in comparison with splitting samples approach: first, we do not need to choose *ad hoc* thresholds (as in Sachs and Warner 1995), and second, by modeling the coefficient itself, we can access how policy variables affect industry productivity growth and convergence.

The remainder of the paper is organized as follows. In the second section, we present data. In the third section, we discuss the underlying specification, motivated by a model *a la* Solow, and make some first exercises with the data. The fourth section presents the estimation method, the fifth section discusses the results, and section six concludes.

4.2 Data

Our industrial database is the same as in Rodrik (2011a, b). We use data from UNIDO's INDSTAT 4 data base, which provides industrial statistics for a wide range of countries at the ISIC 4-digit level (UNIDO 2011). These statistics cover a series of variables, including value added and employment, for up to 127 manufacturing industries per country. As in Rodrik (2011a, b), because of data availability we take 1990 as the starting point for the empirical work. To maximize the number of observations, we estimate pooled regressions using rolling 10-year distances (with 1990 as the starting date) for each industry data available. Our sample includes 127 industries, 49 countries and 8 periods (total of 13,296 observations).

Our educational indicator is years of schooling of the population over the age of 25 (the same as in Glaeser et al. [2004]). Data is from Barro and Lee's Education Attainment Dataset (2011). These variables are provided in 5-year intervals and the gaps are replaced by linear interpolation. We also use the indicator of executive constraint (from Polity IV Data Series version 2010) and the economic openness indicator from the Penn World Table 7.0. These variables are both provided in 1-year intervals. The indicator of executive constraint ranges from 1 (unlimited authority) to 7 (executive parity or subordination). The openness indicator is trade (exports plus imports) as a ratio of GDP.

We also use the vector of latitude and longitude of country's capital as a geographical indicator. Latitude and longitude are proxies for initial endowments, climate and exposition to natural disasters. Data is from the website <http://www.newstrackindia.com>.

The use of variables not from UNIDO's INDSTAT 4 data base reduces the number of countries in our data set because these variables are missing for some countries. Trade openness reduces the number of countries to 48, and the number of observations to 13,265; the executive constraints indicator reduces the number of countries to 38, and the number of observations to 11,363; the years of schooling

indicator reduces the number of countries to 43, and the number of observations to 12,499. Finally, the use of all the three indicators reduces the number of countries to 37, and the number of observations to 11,098. A list of countries in each group is presented in the Appendix J.

4.3 The Underlying Specification and a First Look at Data

4.3.1 The Underlying Solow Model

Our starting point is a textbook Solow model featuring the Cobb-Douglas production function with labor-augmenting technological progress:

$$X_t = K_t^\alpha (A_t L_t)^{1-\alpha}, \quad (4-1)$$

where X_t is output, K_t is capital, and $A_t L_t$ is effective unit of worker. L and A are assumed to grow exogenously at rates n and g so that

$$L_t = L_0 e^{nt}, \quad \text{and} \quad A_t = A_0 e^{gt}. \quad (4-2)$$

Assuming that s is the fraction of output that is saved and invested, and defining output and stock of capital per unit of effective labor as \hat{x} and \hat{k} , respectively, the dynamic equation for \hat{k} is given by

$$\dot{\hat{k}}_t = s \hat{k}_t^\alpha - (n + g + \delta) \hat{k}_t,$$

where δ is the rate of depreciation. \hat{k} converges to its steady state value:

$$\hat{k}^* = \left(\frac{s}{n + g + \delta} \right)^{\frac{1}{1-\alpha}}. \quad (4-3)$$

Substituting equations (4-3) and (4-2) in the log of equation (4-1), gives

$$\ln \hat{x} = \frac{\alpha}{1-\alpha} \ln(s) - \frac{\alpha}{1-\alpha} \ln(n + g + \delta). \quad (4-4)$$

Let \hat{x}^* be the steady state level of income per effective worker. The convergence equation is given by (Islam [1995]):

$$\ln \dot{\hat{x}}_t = \lambda (\ln \hat{x}^* - \ln \hat{x}_t),$$

where $\lambda = (1 - \alpha)(n + g + \delta)$. This equation implies that:

$$\ln \hat{x}_{t_2} - \ln \hat{x}_{t_1} = (1 - e^{-\lambda\tau})(\ln \hat{x}^* - \ln \hat{x}_{t_1}),$$

where $\tau = t_2 - t_1$. Recalling from equation 4-1 that $\hat{x}^* = \hat{k}^{*\alpha}$ and substituting equation (4-3) and $\ln \hat{x}_t = \ln x_t - \ln A_0 - gt$, where x is the income per capita, we

have (Islam [1995]):

$$\begin{aligned} \ln x_{t_2} - \ln x_{t_1} = & - (1 - e^{-\lambda\tau}) \ln x_{t_1} \\ & (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln s \\ & - (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln(n + g + \delta) \\ & + (1 - e^{-\lambda\tau}) \ln A_0 + g(t_2 - e^{-\lambda\tau} t_1), \end{aligned} \quad (4-5)$$

where s is the saving rate and A_0 is the initial level of the labor-augmenting technology.

4.3.2 Industry Productivity Convergence Equation

Call v_{ijt} the log of nominal labor productivity (nominal value added per employee) in industry i , country j and year t . The rate of growth of labor productivity in real terms, Δy_{ijt} , is given by $\Delta y_{ijt} = \Delta v_{ijt} + \pi_{ijt}$, where π_{ijt} is the increase in the industry-level deflator and the Δ before a variable denotes percent changes. Neoclassical growth equations are designed for country aggregates (GDP per capita, country's savings, population growth, among others). To undertake the task of estimating industrial productivity growth, we need an adaptation of equation (4-5). Rodrik (2011b) assumed that the growth in labor productivity in industry is a function of the gap between industry's productivity and its potential (the frontier technology), so

$$\Delta y_{ij,t+1} = \beta(y_{ijt} - y_{it}^*) + D_j, \quad (4-6)$$

where Δy_{ijt} is the growth in the log of labor productivity (measured in US dollars) over some period and D_j is a dummy variable that stands for all time- and industry-invariant country-specific factors. It is a simple adaptation for the industrial sector of equation (4-5). The convergence (or growth) coefficient we are interested in is β .

Assuming a common global U.S. dollar inflation for each individual industry, $\pi_{ijt} = \pi_{ij} + \varepsilon_{ijt}$, and that dollar inflation rates are not systematically correlated with an industry's distance from the technological frontier allow us to express the growth of nominal labor productivity as follows:

$$\Delta v_{ij,t+1} = \beta y_{ijt} + (\pi_{it} - \beta y_{it}^*) + D_j + \varepsilon_{ijt}.$$

Re-arranging terms, we have the following estimating equation

$$\Delta v_{ij,t+1} = \beta y_{ijt} + D_{it} + D_j + \varepsilon_{ij,t+1}, \quad (4-7)$$

where D_{it} is a set of industry and period dummies.¹ The more negative β within a subgroup of countries, the stronger the estimated convergence among them. Or, the larger β , the larger the estimated productivity growth given its initial level. Rodrik (2011b) estimated different versions of equation (4-7). Rodrik (2011b, p.8) argues that a test for unconditional convergence consists of estimating this equation with no country dummies and check whether the estimated convergence coefficient is negative and statistically significant. Tables (4.1) and (4.2) shows the results. In Table (4.2), we weight each observation ijt by the inverse of the probability of country j is sampled, so each country is equally represented. Note in both tables that the convergence coefficient is negative and statistically significant in all specifications with no country dummies (columns (1) to (4)) and the result hardly changes with the inclusion of period and industry dummies. The estimated convergence coefficient seems to be even stronger in the weighted regressions (in the case with period and industry dummies, it is -0.023 in the non-weighted specification and -0.030 in the weighted specification).

¹As noted in Bernard and Durlauf (1996), convergence coefficient estimated in equations like (4-10) may be biased, because it does not include the steady-state level of output (y^*). It is a caveat in our paper as well as in Rodrik (2011a, b).

Table 4.1: Pooled Regressions - 10 year growth rates - 1990-2000 to 1997-2007

	(1)	(2)	(3)	(4)	(5)
$\ln(va/employee)_{t-10}$	-0.020 (45.61)***	-0.019 (45.14)***	-0.023 (52.11)***	-0.022 (50.22)***	-0.066 (80.18)***
Constant	0.239 (52.74)***	0.210 (42.86)***	0.290 (18.10)***	0.199 (3.47)***	0.589 (11.81)***
Period Dummies	No	Yes	Yes	Yes	Yes
Industry Dummies	No	No	Yes	Yes	Yes
Period x Industry Dummies	No	No	No	Yes	Yes
Country Dummies	No	No	No	No	Yes
Observations	13,296	13,296	13,296	13,296	13,296
R-squared	0.14	0.19	0.25	0.29	0.59

Dependent Variable: growth rate of productivity over relevant period

Absolute value of t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.2: Pooled Regressions - 10 year growth rates - 1990 to 2007 - weighted (all countries with the same weight)

	(1)	(2)	(3)	(4)	(5)
$\ln(va/employee)_{t-10}$	(9) -0.024*** (0.00)	(10) -0.023*** (0.00)	(11) -0.030*** (0.00)	(13) -0.028*** (0.00)	(14) -0.073*** (0.00)
Constant	0.266*** (0.02)	0.243*** (0.02)	0.392*** (0.04)	0.318*** (0.02)	0.560*** (0.03)
Period Dummies	No	Yes	Yes	Yes	Yes
Industry Dummies	No	No	Yes	Yes	Yes
Period x Industry Dummies	No	No	No	Yes	Yes
Country Dummies	No	No	No	No	Yes
Observations	13,296	13,296	13,296	13,296	13,296
R-squared	0.12	0.17	0.28	0.39	0.68

Absolute value of t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Dependent Variable: growth rate of productivity over relevant period

We are interested in testing the existence of multiple regime dynamics in industrial productivity growth across countries. So, the whole point in this paper is to allow the convergence β to vary. We estimate the following equation:

$$\Delta v_{ij,t+1} = \beta_j y_{ij,t} + D_i + D_t + \varepsilon_{ijt}. \quad (4-8)$$

Note that we did not include country dummies. This way, our results are directly comparable to the findings in Rodrik (2011b). To reduce the computational cost especially in the semi-parametric specifications, from now on we give up using the interaction of industry and period dummies.²

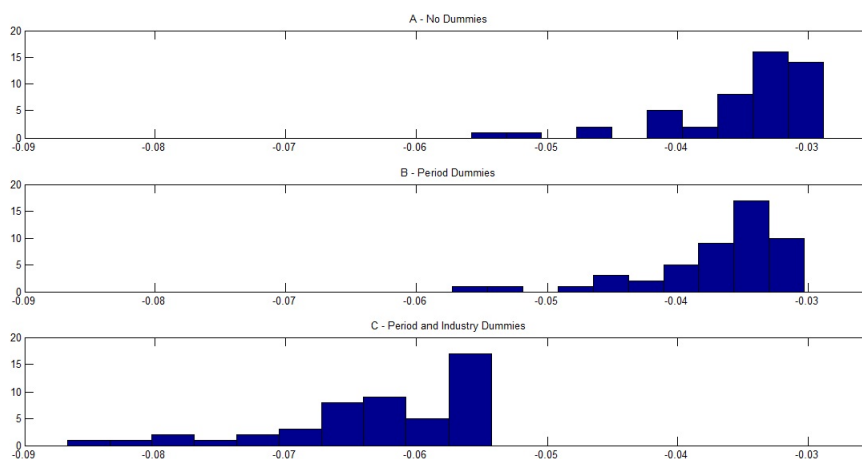
Figure 4.1 shows the histogram of the estimated $\hat{\beta}_j$'s. In panel A, we see the results of regression (4-8) with no dummies; in panel B, period dummies are included; finally, in panel C the equation has industry and period dummies. The $\hat{\beta}_j$'s histograms suggest that the dispersion of the convergence coefficient distribution should not be neglected. For the specification with no dummies, the standard deviation/mean ratio of the $\hat{\beta}_j$'s is 16.2%; in the case with only period dummies, this ratio is 15.3% and, with period and industries dummies, it is 12.3%. We performed Wald tests, in which the null hypothesis is that all countries have the same coefficient. In all three cases (models with no dummies, with only period dummies and with period and industry dummies) the null is rejected (F-statistics around 5000 for the first two cases, and of over 9000 for the last specification)³ Note also that in all the three cases, the estimated convergence coefficients are larger (in absolute value) than the analog estimated coefficients in equation (4-7), shown in table (4.1). Actually, they are much closer to the ones estimated in equation (4-7) where country dummies are included.

In a cross-country regression, the fact that the estimated $\hat{\beta}_j$ is typically negative derives from the empirically suggested fact that industry productivity countries with low industry productivity levels grow faster than the analog for countries with high industry productivity levels. This could be a sign unconditional convergence, i.e., that there is only one steady state level of industry i productivity across countries. But note that if β_j is different across countries, their steady-state levels of productivity are also different. To see that, consider a model with only one industry.⁴ So, equation (4-8)

²We do not believe that it should be a source of concern. Tables (4.1) and (4.2) suggest that, if we are already controlling for industry and period dummies, controlling for the interaction of these dummies causes a very small change in the point estimation and the standard errors of the convergence coefficient.

³We should expect that the distribution of $\hat{\beta}_j$'s is more concentrated the specification with period and industries dummies if part of the difference in the convergence coefficient across countries is due to the production structure.

⁴Multi-industry analysis is similar, we should just condition on industry dummies.

Figure 4.1: β_j 's Histograms

$$\Delta v_{j,t+1} = \alpha_{0t} + \beta_j y_{jt} + \varepsilon_{jt},$$

where $\alpha_{0t} = D_t + \pi_t$. In the steady-state, period effects vanish (i.e., $D_t = D$ and $\pi_t = \pi$), $\Delta v_{j,t+1} = \pi$ and y_{jt} equals the steady-state level y_j^* . We can then write

$$y_j^* = \frac{\pi - \alpha_0}{\beta_j}. \quad (4-9)$$

Therefore, if we can find variables that help us to group countries with the same β_j , we will also be identifying convergence clubs. To capture more accurately the relationship between the relative productivity growth β and country-specific indicators, we allow the convergence coefficient to also vary across decades. This way, we gain one more source of variation. We now estimate bi OLS the following equation:

$$\Delta v_{ij,t+1} = \beta_{jt} y_{ij,t} + D_i + D_t + \varepsilon_{ijt}. \quad (4-10)$$

The exact way this equation is estimated is shown in the Appendix J.

Table (4.3) shows the results of the linear regression of $\hat{\beta}_{jt}$'s (estimated in the equation with industry and period dummies) on various indicators, measured as its decanal initial level. We estimate:

$$\hat{\beta}_{jt} = \gamma_0 + \Gamma' \text{INDIC}_t + \zeta_t, \quad (4-11)$$

where **INDIC** is a combination the following country indicators: latitude, longitude, trade openness, executive constraints and years of schooling. Eight overlapping different decades are covered (1900-2000 through 1997-2007) so that each country enters the data (a maximum of) eight times.⁵

⁵Note that standard errors reported in table (4.3) do not take into account the variance of the

Table 4.3: $\hat{\beta}_{jt}$'s and selected variables in the first year of the decade (t_0). Estimated coefficients are multiplied by 100.

	(1)	(2)	(3)	(4)
Latitude	0.569 (7.50)***	0.241 (3.14)***	0.109 (1.25)	0.103 (1.18)
Longitude	0.010 (0.15)	0.095 (1.75)*	0.048 (0.86)	0.073 (1.43)
Openness, t_0	0.084 (2.02)**			0.061 (1.42)
Executive Constraint, t_0		0.201 (2.60)**		0.168 (1.49)
Years of Schooling, t_0			0.237 (2.81)***	0.328 (3.76)***
Constant	-6.221 (85.59)***	-5.889 (94.68)***	-5.924 (83.83)***	-5.991 (70.63)***
Observations	152	121	135	116
R-squared	0.28	0.24	0.14	0.34

Absolute value of t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Linear regressions indicate that a more educated population in the beginning of the period is associated with a large industry productivity growth. One standard deviation of years of schooling are related to and increase in the convergence coefficient of 0.002 to 0.003, depending on the covariates considered. Once the standard deviation of $\hat{\beta}_{jt}$ is 0.0069 for the model with industry and period dummies, the magnitude of the estimated relation between years of schooling and the convergence coefficient is relevant. Note also that the coefficients of adjustment R^2 of regressions that involve years of education are larger than the ones that do not involve this education indicator.

The relationship between β and trade openness and executive constraints seems to be of less importance. The positive (even though not always statistically significant) coefficient on trade openness indicates that countries with large participation of international trade as a ratio of GDP also faced a large relative productivity growth. This result is quite intuitive, and is in line with studies that relate trade liberalization with productivity gains in industry (for instance, Pavcnik 2002 for the Chilean case and Tybout 2000 for the Mexican case).

The estimated relationship between the executive constraint and the relative productivity growth, although positive when we take only this variable as regressor [column (2) in Table 4.3], is not always statistically significant when trade openness and years of education are also taken into consideration. This result is in line with the findings in Glaeser et al. (2004). The main goal of their work is to assess whether political institutions cause economic growth and the results indicate that poor countries get out of poverty through good policies, often pursued by dictators.

estimated β_{jt} 's from the first step equation 4-10.

4.4 Estimation Method

The industrial convergence equation we are interested in is

$$\Delta v_{ij,t+1} = \beta y_{ijt} + D_i + D_t + \varepsilon_{ijt}.$$

We have presented evidence, however, that the convergence coefficient changes across countries. We have also presented evidence that the time and country-variant convergence coefficient is correlated with some variables. So, the equation we are interested in changes to equation (4-10)

$$\Delta v_{ij,t+1} = \beta_{jt} y_{ijt} + D_i + D_t + \varepsilon_{ijt},$$

where

$$\beta_{jt} = \lambda_0 + f(\mathbf{z}_{jt}; \eta). \quad (4-12)$$

In the above equation, \mathbf{z}_{jt} is a q -dimensional vector of institutional and policy country-specific variables, η is a vector of parameters of limited dimension and $f(\cdot; \cdot)$ is an unknown function. So, the model can be rewritten as

$$\Delta v_{ij,t+1} = (\lambda_0 + f(\mathbf{z}_{jt}; \eta)) y_{ijt} + D_i + D_t + \varepsilon_{ijt}. \quad (4-13)$$

Because there is no economic theory linking these variables to the country and time -specific speed of convergence β_{jt} , we advocate here the use of Neural Network (NN) models to approximate the unknown function in equation (4-13) by

$$t(\mathbf{z}_j; \psi_N) = \left\{ \sum_{m=1}^{M_N} \lambda_m G(\mathbf{z}_j; \boldsymbol{\omega}_m, c_m) \right\}, \quad (4-14)$$

where $\psi_N = (\lambda'_N, \eta'_N)' \in \mathbb{R}^{1+M_N*(2+q)}$, $\lambda_N = \{\lambda_m\}_{m=0}^{M_N} \in \mathbb{R}^{M_N+1}$, $\eta_N = \{\boldsymbol{\omega}_m, c_m\}_{m=1}^{M_N} \in \mathbb{R}^{M_N*(1+q)}$, $\boldsymbol{\omega}_m = (\omega_{1m}, \dots, \omega_{qm})' \in \mathbb{R}^q$ and $c_m \in \mathbb{R}$, $m = 1, \dots, M_N$ are parameters to be estimated,

$$G(\mathbf{z}_j; \boldsymbol{\omega}_m, c_m) = \frac{1}{1 + e^{-(\boldsymbol{\omega}'_m \mathbf{z}_j - c_m)}}, \quad (4-15)$$

and $|\alpha_0| < \infty$, $|\alpha_m| < \infty$, $|c_m| < \infty$, and $\|\boldsymbol{\omega}_m\| < \infty$, $m = 1, \dots, M_N$. The function $G(\cdot; \cdot)$ is called the activation function. Chen et al. (2001) discuss other other choices of activation functions.

As noted in McAleer, Medeiros and Slottje (2008) and Medeiros et al. (2008), most of the recent applied papers concerning NN models havev advocated the “black-box” nature of such kind of specifications, claiming that, due to their “universal approximation” capability, NN models are very flexible and are able to approximate very accurately a vast number of nonlinear mappings. In fact, NNs may be viewed as a kind of smooth transition regression (van Dijk et al., 2002), where the transition variable is an unknown linear combination of the explanatory variables.

In this case there is an optimal number of hidden units, M , that can be translated as the number of limiting regimes (M is fixed) (see, for example, Trapletti et al. (2000), Medeiros and Veiga (2000, 2005)), Medeiros et al. (2006), and Medeiros et al. (2008) for similar interpretations).

On the other hand, when M is large enough, the NN model is an “universal approximator” to any Borel-measurable function over a compact set, and a nonparametric interpretation should be advocated. The number of hidden units increases with the sample size and NN models can be seen as a sieve-approximator of Grenander (1981). Hornik et al. (1994), Chen and Shen (1998), and Chen and White (1998) provide the technical details.

$\psi_N = (\lambda_N, \eta_N) \in \mathbb{R}^r$, $r = 1 + M_N * (q + 2)$, is the vector of all parameters of the model in equation 4 – 13. We advocate the parametric estimation of NN models by making use of the following assumption about the data generating mechanism:

Assumption 6. *There exists a finite constant $M_o \in \mathbb{N}$ and a unique set of parameters $\psi_o = (\lambda_0, \dots, \lambda_{M_o}, \eta_o)$ such that*

$$\Delta v_{ij,t+1} = \beta_{jt} y_{ij,t} + D_i + D_t + \varepsilon_{ijt} = (\lambda_0 + f(\mathbf{z}_{jt}; \eta_o)) y_{ij,t} + D_i + D_t + \varepsilon_{ijt},$$

where

$$f(\mathbf{z}_j; \psi_o) = \left\{ \sum_{m=1}^{M_o} \lambda_m G(\mathbf{z}_j; \boldsymbol{\omega}_m, c_m) \right\}. \quad (4-16)$$

Under Assumption 6, if $\mathbb{E}[\varepsilon_{ijt} | \mathbf{z}_{jt}, y_{ij,t}, D_i, D_t] = 0$, there exist a NN model that can actually correctly approximate the true model when the number of observations goes to infinity. In this case, quasi-maximum likelihood estimators (QMLE) delivery consistent estimators for ψ . The “true” vector of parameters ψ depends on the number of logistic terms M . When NN models are interpreted as semi-parametric devices, M must grow with the sample size. Here, we suppose that there exists one finite numbers M_o such that the “true” data generating mechanism can be approximated arbitrarily well (see McAleer, Medeiros and Slottje 2008).

Assumption 7. *The $((r) \times 1)$ parameter vector ψ_o is an interior point of the compact parameter space Ψ which is a subspace of $\mathbb{R}^r \times \mathbb{R}^1$, the r -dimensional Euclidean space.*

Assumption 8. *The parameters satisfy the conditions $c_1 \leq \dots \leq c_{M_N}$, and $\omega_{qm} > 0 \forall q$ and m .*

Assumption 9. *The model given by equations (4 – 13) to (4 – 15) has no irrelevant hidden units.*

Assumptions 8 and 9 guarantees the global identifiability of the model.

Call N the number of countries and industry units (i.e., there are N combinations of i and j , so that we can refer to the pair (i, j) as the unit n , $n = 1, \dots, N$). Call T the number of fixed time periods. Each unit n is observed a maximum of eight times, from 2000 to 2007. Let the quasi-likelihood function evaluated at an arbitrary parameter ψ be given as

$$\mathcal{L}_N(\psi) = \frac{1}{N} \sum_{n=1}^N \ell_n(\psi) = \frac{1}{N} \sum_{n=1}^N \left\{ -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_n) - \frac{1}{2} \frac{\varepsilon_n^2}{\sigma_n} \right\}.$$

Define the QMLE as

$$\hat{\psi}_N = \arg \max_{\psi \in \Psi} \mathcal{L}_N(\psi) = \arg \max_{\psi \in \Psi} \frac{1}{N} \sum_{n=1}^N \ell_n(\psi).$$

In order to establish the asymptotic normality result, we introduce the following matrices:

$$\mathbf{A}(\psi_o) = \mathbb{E} \left[-\frac{\partial^2 \ell_n(\psi)}{\partial \psi \partial \psi'} \Big|_{\psi_o} \right], \quad \mathbf{B}(\psi_o) = \mathbb{E} \left[\frac{\partial \ell_n(\psi)}{\partial \psi} \frac{\partial \ell_n(\psi)}{\partial \psi'} \Big|_{\psi_o} \right].$$

Proposition 1. *Under Assumptions 1 to 4, the quasi-maximum likelihood estimator $\hat{\psi}$ is almost surely consistent for ψ_o and*

$$N^{1/2}(\hat{\psi}_N - \psi_o) \xrightarrow{d} N(0, \Omega_o), \quad (4-17)$$

where $\Omega_o = \mathbf{A}(\psi_o)^{-1} \mathbf{B}(\psi_o) \mathbf{A}(\psi_o)^{-1}$.

Proof: The strategy of the proof is similar to the ones in McAleer, Medeiros and Slottje (2008) and Medeiros et al. (2008), making appropriate adaptations to our model. To prove consistency we use White (1994), Theorem 3.5, showing that the assumptions stated therein are fulfilled. Assumptions 2.1 and 2.3 in White (1994), related to the probability space and to the density functions, are trivially fulfilled under our assumption (7). Assumption 3.1 in White (1994), related to the existence and continuity of $E[\ell_n(\psi)]$ and is satisfied by our assumption (1). Assumption 3.2 in White (1994) is satisfied by our assumptions (3) and (4), and can also be shown to be related to the negative definiteness of $B_N(\psi)$.

To prove asymptotically normality of the QMLE we follow McAleer, Medeiros and Slottje (2008), Theorem (2). We need the following conditions in addition to those to proof consistency (see also White (1994, page 92)): (1) The true parameter vector ψ_o is interior to Ψ ; (2) the matrix $\mathbf{A}_N(\psi)$ exists and is continuous in Ψ ; (3) $\mathbf{A}_N(\psi) \xrightarrow{p} \mathbf{A}(\psi)$ for all ψ in Ψ ; (4) the score vector satisfies

$$\frac{1}{N} \sum_{i=1}^N \left[\frac{\partial \ell_i(\psi)}{\partial \psi} \Big|_{\psi_o} \right] \xrightarrow{d} N[\mathbf{0}, \mathbf{B}(\psi_o)], \quad (4-18)$$

where $\mathbf{B}(\psi_o)$ is positive definite.

Condition (1) is satisfied by our assumption (7). Condition (2) follows from the fact that ℓ_i is differentiable of order two on $\psi \in \Psi$ under assumption (1). Condition (3) is also trivial to verify. Finally, standard application of the Central Limit Theorem guarantees that Condition (8) holds.

In applications, the number of sieves (or hidden units) is unknown and should be determined from the data. In the neural network literature several approaches have been proposed. A popular method for doing that is pruning, in which a model with a large number of hidden units is estimated first, and the size of the model is subsequently reduced by applying an appropriate technique such as cross-validation. Another technique used in this connection is regularization, which may be characterized as penalized maximum likelihood or least squares applied to the estimation of neural network models. Bayesian regularization, based on selecting a prior distribution for the parameters, may serve as an example.

Another possibility, which is adopted in the paper, is to estimate R models, with $M_N = 0, \dots, R$, for R sufficiently large, and choose the optimal M_N^* based on the use of model selection criteria (MSC) or cross-validation. In this paper we fix $R = 5$ and advocate the use of the Schwarz Bayesian Information Criterion (SBIC) to determine the number of sieves.

4.5 Results

Our main target is to explain the differences in β_{jt} defined in equation (4-8) by institutional and policy variables. Because there is no economic theory linking these variables to the relative productivity growth β_{jt} , we use NN models, taking advantage of their “universal approximation” capability.

Table 4.4 presents the country-and-year mean partial effects of latitude, longitude, trade openness, years of schooling and executive constraints on the relative productivity growth. These values are the partial effects of one standard deviation increase in these variables on the estimated convergence coefficient. Under Assumption 6, the partial effect of variable z_q on convergence coefficient is

$$\frac{\partial \beta_{jt}}{\partial z_{jq}}(\mathbf{z}_j; \psi) = \sum_{m=1}^{M_N} \lambda_m \omega_{qm} [G(\mathbf{z}_j; \boldsymbol{\omega}_m, c_m) (1 - G(\mathbf{z}_j; \boldsymbol{\omega}_m, c_m))]$$

We can estimate the parameters that govern the distribution of parameters ψ (equation 4-17). It allows us do Monte Carlo experiments and compute means and standard deviations of the mean partial effects, which are reported in Table 4.4. Table 4.4 is the semi-parametric analog of Table 4.3. Indeed, note that the point estimates are similar. In the more complete specifications in column (4), Table 4.3, the estimated mean partial effects of trade openness, executive constraint and years of schooling are, respectively, 0.061, 0.168 and 0.328; in Table 4.4, the corresponding mean partial effects are 0.072, 0.147, 0.182. In our more complete specification in equation (4), trade openness, executive constraints and years of schooling are statistical significant. However, as opposed to years of schooling, the statistical significance of the partial effect of latitude, longitude, trade openness and executive constraints does not seem to be robust. Depending on which variables are considered in the model, and which dummies are included, the standard deviation of the point estimates of the mean partial effects of these variables can be high.⁶

Figures 4.2 and 4.3 help us understand this. Let us first focus on Figure 4.2. Each panel in Figure 4.2 plots the mean (across Monte Carlo simulations) of the mean (across observed countries and years) values of β'_{jt} s as a function of observed values of normalized (zero mean, variance one) values of trade openness (Panel A), executive constraints (Panel B) and years of schooling (Panel C) measured by the model with latitude, longitude, openness, executive constraints, years of schooling and period and industry dummies. In panel A, for example, we change the values trade openness along the observed range and check how β_{jt} changes, keeping all other variables fixed in their sample mean.

Panel A indicates that, up to a certain point, opening the economy to foreign trade makes the relative productivity growth increase. But results suggest that, when trade openness is high, everything else constant, additional opening reduces growth. The statistical non-significance of trade openness in Table 4.4 may be explained by this non-linear behavior. The effect of executive constraint is always positive and almost linear, indicating that, for this variable, our semi-parametric tool does not add much information to the analysis. Panel C in Figure 4.2 suggests that the effect of years of schooling has an interesting concave shape. For countries with low levels of education, increases in years of schooling brings large gains in growth, but gains diminish as education improves, reaching negative values for countries with high levels of education. Panel C brings evidence that convergence is weaker in within groups with higher levels of education. This finding parallels the results of Durlauf and Johnson (1995) who failed to find evidence of convergence among the high-output economies, and De Long (1988) who rejected convergence over a much

⁶See in Appendix L the versions of tables 4.4 and 4.3 in exercises where the convergence coefficient is estimated in an equation with no dummies and with only period dummies.

Table 4.4: Country-and-year mean partial effects
 . Industry productivity growth regressions: 10-year growth rates.
 Dependent variable: growth rate of labor productivity over relevant period.
 Estimated coefficients and standard values are multiplied by 100.

	(1)	(2)	(3)	(4)
Latitude	0.247 (0.04)***	0.172 (0.35)	-0.033 (0.06)	0.070 (0.05)
Longitude	-0.026 (0.02)	0.095 (0.22)	-0.080 (0.10)	-0.035 (0.05)
Openness	-0.020 (0.07)			0.072 (0.02)***
Executive Constraint		0.319 (0.22)		0.147 (0.07)**
Years of Schooling			0.263 (0.04)***	0.182 (0.08)**
Number of Sieves	5	5	5	5
Number of Observations	13,265	11,363	12,499	11,098
R-squared (underlying regression)	0.34	0.38	0.42	0.41

Standard deviations in parenthesis based on Monte Carlo simulations (1000 repetitions)

longer time span (from 1870 to 1979) when studying economies with similar high initial outputs.

Figure 4.3 plots the partial effect itself, i.e., the slope of curves presented in Figure 4.2. The continuous line is the estimated partial effect of variables in the semi-parametric model. The dashed line is the mean (across Monte Carlo simulations) of the mean (across observed countries and years) partial effect of variables in the semi-parametric model (the same value as in Table 4.4). The dotted line is the partial effect as measured in the linear model (the same value as in Table 4.3). Observing Figure 4.3 helps us understand what we loose estimating the productivity growth equation linearly. Note that, specially for openness and years of schooling, the partial effect of variables changes importantly as the level the variable that originates the effect changes.

Finally, figure 4.4 shows the estimated convergence coefficient median (central mark) and 90% confidence interval (box) by country (period mean) as measured by the model with latitude, longitude, years of schooling, and period and industry dummies. This figure can be seen as an exercise to group countries with the same convergence coefficient. We discussed in subsection 4.3.2, equation (4-9), that if we can find variables that help us to group countries with the same β_j , we will also be identifying convergence clubs. Let us use the United States as an example. The edges of the confidence interval box for this country are marked with light dashed lines. With 90% of confidence, we can not reject that the United States is in the same group as Italy, France, until Israel, following the order of countries in Figure 4.4. But, we reject that the Unites States is in the group of Thailand, Mauritius, until

Figure 4.2: β_{jt} and normalized (zero mean, variance one) values of openness (Panel A), executive constraints (Panel B) and years of schooling (Panel C) measured by the model with latitude, longitude, openness, executive constraints, years of schooling and period and industry dummies.

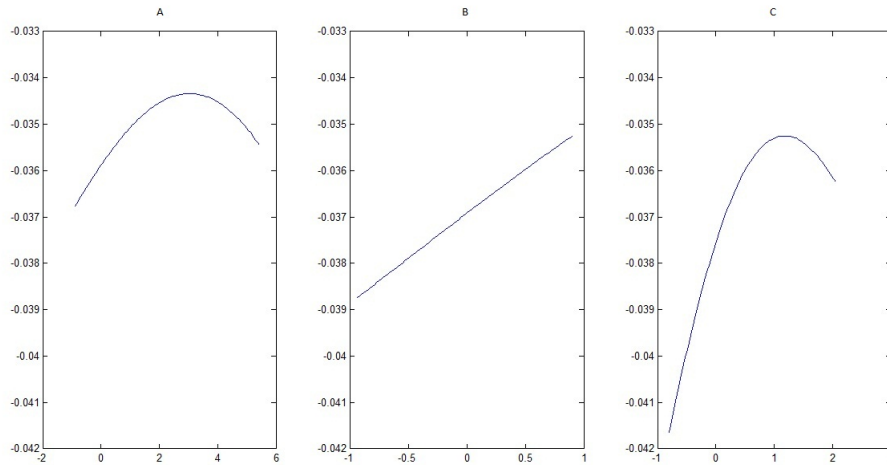


Figure 4.3: Partial effects and normalized (zero mean, variance one) values of openness (Panel A), executive constraints (Panel B) and years of schooling (Panel C) measured by the model with latitude, longitude, openness, executive constraints, years of schooling and period and industry dummies. The dashed line is the partial effect of variables in the semi-parametric model and the dotted line is the partial effect as measured in the linear model.

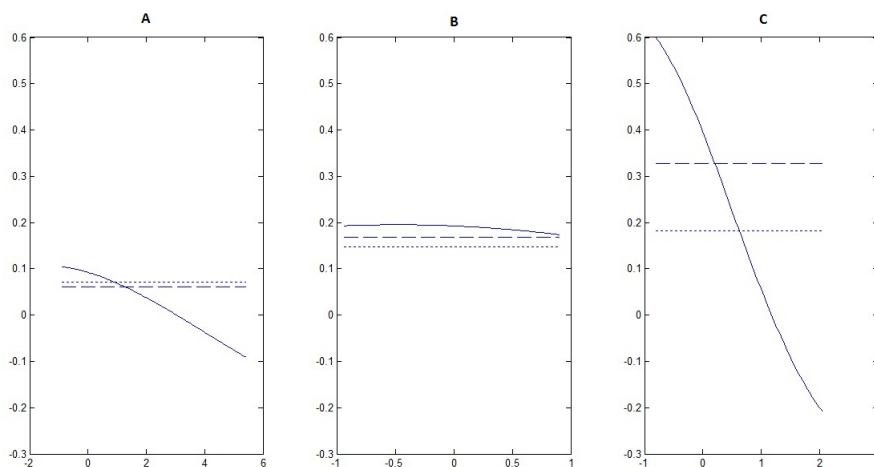
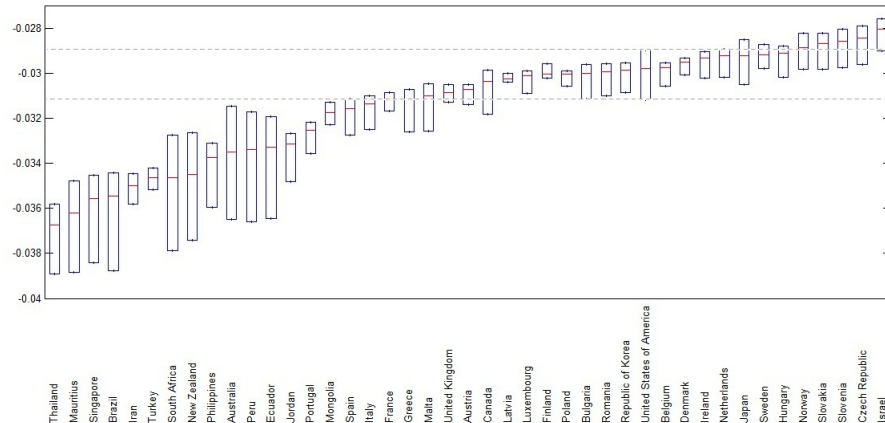


Figure 4.4: Estimated convergence coefficient median (central mark) and 90% confidence interval (box) by country (period mean) as measured by the model with latitude, longitude and years of schooling. Models with period and industry dummies.



Spain, following the order of countries.

The differences in the convergence coefficient across countries is not only of statistical significance, but it is also economically meaningful. The extreme values of the estimated convergence coefficient shown in figure 4.4 are -3.7% per year for Thailand, and -2.8% per year for Israel. It means that the half life to productivity convergence for Thailand is of 27 years and, for Israel, it is of 37 years.

Here, we present is a different way of grouping countries in convergence clubs. In this way, country-specific variables tell us how countries must be grouped, in some confidence interval, and we can also estimate how these variables are related to the convergence coefficient.

4.6 Concluding Remarks

Our goal in this paper were not to access how institutions affect industry convergence. We are less ambitious. In the first place, the word “affect” implies causality and we are not denying the well established fact that a country’s growth (and specifically, its industrial growth) over time affects its institutional variables [see Acemoglu et al. (2001) and Glaeser et al. (2004)]. Secondly, as noted in Glaeser et al. (2004), the word “institution” is related to constraints on individual behavior and need to be reasonably permanent or durable. The fact the we use variables such as educational and trade openness indicators makes clear that we are not pursuing *permanent constraints*. After the work of Rodrik (2011a, b), that helped us

to understand that we should search for unconditional convergence in the industry sector, our goal is to establish two related facts : (i) The pace of convergence is different across countries; and (ii) The pace of convergence is related to some important variables.

Bibliography

Acemoglu, Daron, Simon Johnson and James A. Robinson, 2001, *The Colonial Origins of Comparative Development: An Empirical Investigation*, American Economic Review 91(5), 1369–1401.

Ai, C. and X. Chen, 2003 *Efficient Estimation of Models with Conditional Moment Restrictions Containing*, Econometrica 71,1795-1843.

Almeida, C. and B. Meres, 2008, *Extracting Default Probabilities from Sovereign Bonds*, Brazilian Review of Econometrics 28, 77-94.

Amato, J. and M. Luisi, 2006, *Macro Factors in the Term Structure of Credit Spreads*, Bank of International Settlements Working Paper, Number 203.

Anderson, B., P. J. Hammond and C. A. Ramezani, 2010, *Affine Models of the Joint Dynamics of Exchange Rates and Interest Rates*, Journal of Financial and Quantitative Analysis, 45(5), 1341-1365.

Andritzky, J., 2006, *Sovereign Default Risk Valuation: Implications of Debt Crises and Bond Restructurings*, Springer Berlin Heidelberg, New York.

Ang, A. and M. Piazzesi, 2003, *A No-arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables*, Journal of Monetary Economics 50, 745–787.

Ang, A., M. Piazzesi, and M. Wei, 2006, *What does the yield curve tell us about GDP growth?*, Journal of Econometrics 131, 359-403.

Arora, V. and M. Cerisola, 2001, *How does U.S. Monetary Policy Influence Sovereign Spreads in Emerging Markets?*, IMF Staff Papers, 48(3), 474-498.

Azariadis, Costas and Allan Drazen, 1990, *Threshold Externalities in Economic Development* The Quarterly Journal of Economics 105(2), 501–526.

Backus, D., S. Foresi and C. Telmer, 2001, *Affine Term Structure Models and the Forward Premium Anomaly*, The Journal of Finance, 56(1), 279-304.

Baker, S. R., N. Bloom and S. J. Davis, 2012, *Measuring Economic Policy Uncertainty*, Working Paper, available at

<http://www.policyuncertainty.com/media/BakerBloomDavis.pdf>.

Baldacci, E. and M. Kumar, 2010, *Public Debt and Growth*, IMF Working Paper, Number 10/174.

Baumol, William J., 1986, *Productivity Growth, Convergence, and Welfare: What the Long-run Data Show*, The American Economic Review 76(5), 1072–85.

Barro and X. Sala-i-Martin, R.J., 1992 *Convergence*, Journal of Political Economy, 100, 223-251.

Bernard, A. B. and S. N. Durlauf, 1996, *Interpreting tests of the convergence hypothesis*, Journal of Econometrics, 71, 161–173.

Bernoth, K., J. von Hagen and L. Schuknecht, 2004, *Sovereign Risk Premiums in the European Government Bond Market*, European Central Bank Working Paper, Number 369.

Bloom, N., 2009, *The Impact of Uncertainty Shocks* Econometrica, 77(3), 623-685.

Bonomo, M. and A. Lowenkron, 2008, *A Term Structure Model for Defaultable Bonds with Macro and Latent Variables*, Working Paper, available at ftp://ftp.econ.puc-rio.br/PDF/seminario/2008/Bonomo_Lowenkron%282007%29.pdf.

Canova, Fabio, 2004, *Testing for Convergence Clubs in Income Per Capita: A Predictive Density Approach*, International Economic Review Volume 45, Issue 1, 49–77.

Chen, R. and R. S. Tsay, 1991 *On the Ergodicity of TAR(1) Processes*, Annals of Applied Probability, 1, 613-634.

Chen, Xiaohong, 2007. *Large Sample Sieve Estimation of Semi-Nonparametric Models,* *Handbook of Econometrics*, in: J.J. Heckman and E.E. Leamer (ed.), *Handbook of Econometrics*, Edition 1, Volume 6, Chapter 76, Elsevier.

Chen, X. and X. Shen, 1998, *Sieve extremum estimates for weakly dependent data*, Econometrica 66, 289–314. Chen, X. and H. White, 1998, *Improved rates and asymptotic normality for nonparametric neural network estimators*, IEEE Transactions on Information Theory 18, 17–39. Cochrane, J. H. and M. Piazzesi, 2008, *Decomposing the Yield Curve*, Working Paper.

Dai, Q. and K. Singleton, 2000, *Specification Analysis of Affine Term Structure Models*, The Journal of Finance, 55(5), 1943-1977.

Davison, A. C. and D. V. Hinkley, 1997, *Bootstrap Methods and Their Application*, 10th edition, Cambridge University Press, New York.

Diebold, F. X., G. Rudebusch and S. B. Aruoba, 2006, *The macroeconomy and the yield curve: a dynamic latent factor approach*, *Journal of Econometrics* 131, 309-338.

De Long, J. B., 1988, *Productivity Growth, Convergence, and Welfare: Comment*, *The American Economic Review*, 78(5), 1138–1154.

de los Rios, A. D., 2009, *Can affine term structure models help us predict exchange rates?*, *Journal of Money, Credit and Banking*, 41(4), 755-766.

Durlauf, Steven N. and Paul A. Johnson, 1995, *Multiple Regimes and Cross-Country Growth Behaviour*, *Journal of Applied Econometrics*, 10(4), 365–84.

Durlauf, Steven N., Paul A. Johnson, and Jonathan R.W. Temple, 2005, *Growth Econometrics*, in: Philippe Aghion and Steven Durlauf (ed.), *Handbook of Economic Growth*, Edition 1, Volume 1, Chapter 8, 555–677, Elsevier North-Holland, Amsterdam.

Duffee, G. R., 1996, *Treasury yields and corporate bond yield spreads: an empirical analysis*, Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series, Volume 96-20.

Duffee, G. R., 1998, *The relation between Treasury yields and corporate bond yield spreads*, *The Journal of Finance*, 53(6), 2225-2242.

Duffee, G. R., 2002 *Term premia and interest rate forecasts in affine models*, *The Journal of Finance*, 57(1), 405-443.

Duffee, G. R., 2011, *Information in (and not in) the Term Structure*, *Review of Financial Studies*, 24(9), 2895-2934.

Duffie, D. and K. Singleton, 1999, *Modeling term structures of defaultable bonds*, *Review of Financial Studies* 12, 687-720.

Duffie, D. and K. Singleton, 2003, *Credit Risk: Pricing, Measurement and Management*, Princeton University Press.

Duffie, D., L. Pedersen and K. Singleton, 2003, *Modeling credit spreads on sovereign debt: A case study of Russian bonds*, *Journal of Finance* 55, 119-159.

Duffie, D. and R. Kan, 1996, *A yield-factor model of interest rates*, *Mathematical Finance* 6, 379-406.

Dufour, J.-M. ,1982 *Generalized Chow Tests for Structural Change: A*

- Coordinate-Free Approach*, *International Economic Review*, 23, 565-575.
- Durre, A. and F. Smets, 2012, *Sovereign debt and monetary policy in theory and practice: The case of the Euro area*, in J. Chadha, A. Durre, M. Joyce and L. Sarno (ed.), *New Developments in Macro-Finance Yield Curves*.
- Eichengreen, Barry, 2011, *Exorbitant Privilege*, Oxford University Press, New York.
- Elmslie, B.T. ,1995 *Retrospectives: the Convergence Debate Between David Hume and Josiah Tucker*, *Journal of Economic Perspective*, 9, 207-216.
- Favero, C., 2002, *How Do European Monetary and Fiscal Authorities Behave?*, CEPR Discussion paper, Number 3426.
- Fisher, F.M., 1970 *Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note*, *Econometrica*, 38,361-366.
- Glaeser, Edward L., Rafael La Porta, Florencio Lopez—de—Silanes, Andrei Shleifer, 2004, *Do Institutions Cause Growth?*, *Journal of Economic Growth* 9(3), 271—303.
- Georgoutsos, D. A. and P. M. Migiakis, 2010, *European sovereign bond spreads: monetary unification, market conditions and financial integration*, Bank of Greece Working Paper, Number 115.
- Grenander, U., 1981, *Abstract Inference*. Wiley Series, New York.
- Hornik, K., M. Stinchcombe, H. White and P. Auer, 1994, *Degree of approximation results for feedforward networks approximating unknown mappings and their derivatives*, *Neural Computation* 6, 1262—1274.
- Horowitz, J. L., 2001, *The Bootstrap*, in J.J. Heckman and E.E. Leamer (ed.), *Handbook of Econometrics Edition 1, Volume 5, Number 5*, Elsevier.
- Hsiao, C., 1989, *Analysis of Panel Data*, Cambridge University Press.
- Hwang, Jason J., 2007, *Patterns of Specialization and Economic Growth*, unpublished Ph.D dissertation, Economics Department, Harvard University.
- Islam, Nazrul, 1995, *Growth Empirics: A Panel Data Approach*, *The Quarterly Journal of Economics* 110(4), 1127—70.
- Islam, Nazrul, 2003, *What have We Learnt from the Convergence Debate?*, *Journal of Economic Surveys* 17(3),309—362.
- Joslin, S., A. Le and K. Singleton,2012, *Why Gaussian Macro-Finance Term*

Structure Models are (Nearly) Unconstrained Factor-VARs, Journal of Financial Economics, forthcoming.

Joslin, S., M. Pribsch, and K. Singleton, 2010, *Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks*, Working Paper.

Litterman, R. and J. Scheinkman, 1991, *Common factors affecting bond returns*, The Journal of Fixed Income, 1(1), 54-61.

Longstaff, F. A., J. Pan, L. H. Pedersen and Kenneth J. Singleton, 2011, *How Sovereign is Sovereign Credit Risk*, American Economic Journal: Macroeconomics 3, 75-103.

Mankiw, N. Gregory, David Romer and David N. Weil, 1992, *A Contribution to the Empirics of Economic Growth*, The Quarterly Journal of Economics 107(2), 407–437.

Marattin, L., P. Paesani and S. Salotti, 2012, *Some considerations on debt and interest rates*, in J. Chadha, A. Durre, M. Joyce and L. Sarno (ed.), *New Developments in Macro-Finance Yield Curves*.

McAleer, Michael, Marcelo C. Medeiros, Daniel Slottje, 2008, *A Neural Network Demand System with Heteroskedastic Errors*, Journal of Econometrics, 147, 359–371.

Medeiros, Marcelo C. and Álvaro Veiga, 2000, *A hybrid linear-neural model for time series forecasting*, IEEE Transactions on Neural Networks 11, 1402–1412.

Medeiros, Marcelo C. and Álvaro Veiga, 2005, *A Flexible Coefficient Smooth Transition Time Series Model*, IEEE Transactions on Neural Networks 16(1), 97–113.

Medeiros, Marcelo C., Michael McAleer, Daniel Slottje, Vicente Ramos and Javier Rey-Maqueira, 2008, *An Alternative Approach to Estimating Demand: Neural Network Regression with Conditional Volatility for High Frequency Air Passenger Arrivals*, Journal of Econometrics, 147, 372–383.

Medeiros, Marcelo C., T. Terasvirta, G. Rech, 2006, *Building neural network models for time series: A statistical approach*. Journal of Forecasting 25, 49–75.

Meres, B. and C. Almeida, 2008, *Extracting Default Probabilities from Sovereign Bonds*, Revista de Econometria, 28, 77-94.

O'Neill, Jim, 2011, *The Growth Map: Economic Oportuninies in the BRICs and Beyond*, Penguin Books, London.

- Pan, J. and K. Singleton, 2008, *Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads*, *Journal of Finance* 63, 2345-23848.
- Pavcnic, N., 2002, *Trade Liberalisation, exit and productivity improvements: Evidence from Chilean plants*, *Review of Economic Studies*, 69, 245–276.
- Perotti, R., 2005, *Estimating the effects of fiscal policy in OECD countries*, Proceedings, Federal Reserve Bank of San Francisco.
- Robinson, P.M., 1988 *Root n Consistent Semiparametric Regression*, *Econometrica*, 56, 931-954.
- Rodrik, Dani, 2011a, *The Future of Economic Convergence*, NBER Working Paper Number 17400.
- Rodrik, Dani, 2011b, *Unconditional Convergence*, NBER Working Paper Number 17546.
- Sachs, Jeffrey and Andrew M. Warner, 1995, *Economic Convergence and Economic Policies*, NBER Working Paper Number 5039.
- Schuknecht, L., J. von Hagen and G. Wolswijk, 2008, *Government Risk Premiums in the Bond Market, EMU and Canada*, European Central Bank Working Paper, Number 879.
- Spence, Michael, 2011, *The Next Convergence: The Future of Economic Growth in a Multispeed World*, Farrar, Straus and Giroux, New York.
- Subramanian, Arvind, 2011 *Eclipse: Living in the Shadow of China's Economic Dominance*, Peterson Institute for International Economics, Washington DC.
- Trapletti, A., F. Leisch, F. and K. Hornik, 2000, *Stationary and integrated autoregressive neural network processes*, *Neural Computation* 12, 2427–2450.
- Tybout, J., 2000, *Manufacturing firms in developing countries: how well do they do and why*, *Journal of Economic Literature* 38(1).
- van Dijk, D., T. Terasvirta and P. H. Franses, 2002, *Smooth transition autoregressive models – A survey of recent developments*, *Econometric Reviews* 21, 1–47.
- White, H., 1994, *Estimation, Inference and Specification Analysis*, Cambridge University Press, New York, NY. Wu, L. and F. X. Zhang, 2008, *A no-arbitrage analysis of macroeconomic determinants of the credit spread term structure* *Management Science*, 54(6), 1160-1175.

A Riccati Difference Equations

The stochastic discount rate M_{t+1} can be written as:

$$M_{t+1} = \exp \left\{ -\frac{1}{2} \lambda_t' \lambda_t - \delta_0 - \delta_1' \mathbf{X}_t - \lambda_t' \varepsilon_{t+1} \right\}, \quad (\text{A-1})$$

where λ_t is the market price of risk (time variant) associated with the sources of uncertainty ε_t . λ_t is parameterized as an affine function of the state variables:

$$\lambda_t = \lambda_0 + \lambda_1' \mathbf{X}_t, \quad (\text{A-2})$$

for a $K \times 1$ vector λ_0 a $K \times K$ matrix λ_1 . From Ang and Piazzesi (2003), we know:

Result 1: If $M_{t+1} V_{t+1}^{(N)}$ follows a lognormal distribution, the restriction of no arbitrage implies that

$$\bar{A}_{N+1} = \bar{A}_N + \bar{\mathbf{B}}_N' (-\Sigma^{\text{Ge}} \lambda_0) + \frac{1}{2} \bar{\mathbf{B}}_N' \Sigma^{\text{Ge}} \Sigma^{\text{Ge}'} \bar{\mathbf{B}}_N - \delta_0^{\text{Ge}} \quad (\text{A-3})$$

$$\bar{\mathbf{B}}_{N+1}' = \bar{\mathbf{B}}_N' \left(\Phi^{\text{Ge}} - \Sigma^{\text{Ge}} \lambda_1' \right) - \delta_1^{\text{Ge}'} \quad (\text{A-4})$$

with $\bar{A}_1 = -\delta_0^{\text{Ge}}$ and $\bar{\mathbf{B}}_1' = -\delta_1^{\text{Ge}'}$. The yield $y_t^{(N)}$ at t of a bond with maturity N is given by:

$$y_t^{(N)\text{Ge}} = -\frac{\log(V_t^{(N)\text{Ge}})}{N} = A_N + \mathbf{B}_N' \mathbf{X}_t^{\text{Ge}}, \quad (\text{A-5})$$

where $A_N = -\frac{\bar{A}_N}{N}$ and $\mathbf{B}_N = -\frac{\bar{\mathbf{B}}_N}{N}$. In short, we need to estimate the parameters $\Psi^{\text{Ge}} = (\delta_0^{\text{Ge}}, \delta_1^{\text{Ge}'}, \Phi^{\text{Ge}}, \Sigma^{\text{Ge}}, \lambda_0, \lambda_1)$. Recall that δ_0^{Ge} is a scalar; $\delta_1^{\text{Ge}'}$ and λ_0 are K -dimensional vectors; $\Phi^{\text{Ge}}, \Sigma^{\text{Ge}}$ and λ_1 are $K \times K$ matrices. Before restrictions are imposed, Ψ^{Ge} can be written as a $(1+2K+3K^2)$ -dimensional vector of parameters.

B Proof of Proposition 1

The proof of Proposition 1 has two steps: In the first one, we prove that, given our assumptions, the short rate spread $s_t^1 = y_t^1 - y_t^{1Ge}$ is linear in \mathbf{X}_t ; in the second step, we show that, given that the short rate spread is linear in \mathbf{X}_t and the dynamics of Germany's yield curve given by Result 1, spreads of other maturities are given by equations (1-9) to (1-11).

The short rate spread $s_t^1 = y_t^1 - y_t^{1Ge}$ is linear in \mathbf{X}_t . Consider V_t^N the price at t of a bond that pays no coupon and makes a payment of \$ 1 at time $t + N$. For $N=0$, $V_t^N = 1$. Hence, equation (1-7) can be written as

$$V_t^1 = E_t [M_{t+1} \Theta_{t+1}^0]. \quad (\text{B-1})$$

If $M_{t+1} \Theta_{t+1}^N V_{t+1}^N$ follows a lognormal distribution, equation (B-1) gives $v_t^1 = E_t [m_{t+1} + \theta_{t+1}^0] + \frac{1}{2} \text{Var}_t [m_{t+1} + \theta_{t+1}^0]$, where lowercase letters denote variables in *log*. Substituting equations (1-8) and (A-1), we have:

$$\begin{aligned} v_t^1 &= E_t \left[-\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} - \lambda_t' \varepsilon_{t+1}^{Ge} + \theta_{00} + 0 * \theta_{01} + \theta_1' (\Phi \mathbf{X}_t + \Sigma \varepsilon_{t+1}) \right] \\ &+ \frac{1}{2} \text{Var}_t \left[-\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} - \lambda_t' \varepsilon_{t+1}^{Ge} + \theta_{00} + 0 * \theta_{01} + \theta_1' (\Phi \mathbf{X}_t + \Sigma \varepsilon_{t+1}) \right] \\ &= E_t \left[-\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + \theta_1' \Phi \mathbf{X}_t \right] + \frac{1}{2} \text{Var}_t \left[-\lambda_t' \varepsilon_{t+1}^{Ge} + \theta_1' \Sigma \varepsilon_{t+1} \right] \\ &= -\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + \theta_1' \Phi \mathbf{X}_t + \frac{1}{2} \text{Var}_t \left[(-\lambda_t' \mathbf{J} + \theta_1' \Sigma) \varepsilon_{t+1} \right], \end{aligned}$$

where \mathbf{J} is a selection ($K^{Ge} \times K$)-dimensional matrix such that $\varepsilon_t^{Ge} = \mathbf{J} \varepsilon_t$. Note that $\mathbf{J} \mathbf{J}' = \mathbf{I}_{K^{Ge}}$. So, $v_t^1 = -\delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + \theta_1' \Phi \mathbf{X}_t - \theta_1' \Sigma \mathbf{J}' \lambda_t + \frac{1}{2} \theta_1' \Sigma \Sigma' \theta_1$. Substituting equation (A-2), $v_t^1 = -\delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + \theta_1' \Phi \mathbf{X}_t - \theta_1' \Sigma \mathbf{J}' (\lambda_0 + \lambda_1' \mathbf{X}_t^{Ge}) + \frac{1}{2} \theta_1' \Sigma \Sigma' \theta_1$. We know from equation (1-1) that $v_t^{1Ge} = -r_t^{Ge} = -\delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge}$. So, $v_t^1 = v_t^{1Ge} + [\theta_{00} - \theta_1' \Sigma \mathbf{J}' \lambda_0 + \frac{1}{2} \theta_1' \Sigma \Sigma' \theta_1] + [\theta_1' (\Phi - \Sigma \mathbf{J}' \lambda_1' \mathbf{J})] \mathbf{X}_t$. It means that

$$v_t^1 = v_t^{1Ge} + \bar{D}_1 + \bar{\mathbf{E}}_1' \mathbf{X}_t, \quad (\text{B-2})$$

where $\bar{D}_1 = [\theta_{00} - \theta_1' \Sigma \mathbf{J}' \lambda_0 + \frac{1}{2} \theta_1' \Sigma \Sigma' \theta_1]$, and $\bar{\mathbf{E}}_1' = [\theta_1' (\Phi - \Sigma \mathbf{J}' \lambda_1' \mathbf{J})]$. Equation (B-2) can be written the following way: $s_t^1 = y_t^1 - y_t^{1Ge} = -\frac{v_t^1}{1} + \frac{v_t^{1Ge}}{1} = D_1 + \mathbf{E}_1' \mathbf{X}_t$, where $D_1 = -\bar{D}_1/1$ and $\mathbf{E}_1 = -\bar{\mathbf{E}}_1/1$.

Second step: given that the short rate spread is linear in X_t and the dynamics of Germany yield curve is given by Result 1, spreads of other maturities are given by equations (1-9) to (1-11). From equation (1-7), and assuming that $M_{t+1} \Theta_{t+1}^N V_{t+1}^N$ follows a lognormal distribution, we have

$$v_t^{N+1} = E_t [m_{t+1} + \theta_{t+1}^N + v_{t+1}^N] + \frac{1}{2} \text{Var}_t [m_{t+1} + \theta_{t+1}^N + v_{t+1}^N].$$

Take $N=1$. Substituting equations (A-1), (A-2), (1-2), (B-2) and (1-8) into the above equation gives:

$$\begin{aligned} v_t^{N+1} = & -\frac{1}{2} \lambda'_t \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + N * \theta_{01} + \theta'_1 \Phi \mathbf{X}_t + \bar{A}_N + \bar{\mathbf{B}}'_N \Phi^{Ge} \mathbf{X}_t^{Ge} \\ & + \bar{D}_N + \bar{\mathbf{E}}'_N \Phi \mathbf{X}_t \\ & + \frac{1}{2} \text{Var}_t \left[(\bar{\mathbf{B}}'_N \Sigma^{Ge} - \lambda'_t) \varepsilon_{t+1}^{Ge} + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \varepsilon_{t+1} \right]. \end{aligned} \quad (\text{B-3})$$

Opening only $\text{Var}_t \left[(\bar{\mathbf{B}}'_N \Sigma^{Ge} - \lambda'_t) \varepsilon_{t+1}^{Ge} + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \varepsilon_{t+1} \right]$, we have:

$$\begin{aligned} & \text{Var}_t \left[(\bar{\mathbf{B}}'_N \Sigma^{Ge} - \lambda'_t) \varepsilon_{t+1}^{Ge} + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \varepsilon_{t+1} \right] \\ & = \text{Var}_t \left\{ \left[(\bar{\mathbf{B}}'_N \Sigma^{Ge} - \lambda'_t) \mathbf{J} + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \right] \varepsilon_{t+1} \right\} \\ & + \lambda'_t \lambda_t - 2 \bar{\mathbf{B}}'_N \Sigma^{Ge} \lambda_t + \bar{\mathbf{B}}'_N \Sigma^{Ge} \Sigma^{Ge'} \bar{\mathbf{B}}_N - 2 \theta'_1 \Sigma \mathbf{J}' \lambda_t + 2 \theta'_1 \Sigma \mathbf{J}' \Sigma^{Ge'} \bar{\mathbf{B}}_N \\ & - 2 \bar{\mathbf{E}}'_N \Sigma \mathbf{J}' \lambda_t + 2 \bar{\mathbf{E}}'_N \Sigma \mathbf{J}' \Sigma^{Ge'} \bar{\mathbf{B}}_N + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \Sigma' (\theta'_1 + \bar{\mathbf{E}}'_N)'. \end{aligned}$$

Substituting into equation (B-3),

$$\begin{aligned} v_t^{N+1} = & -\frac{1}{2} \lambda'_t \lambda_t - \delta_0^{Ge} - \delta_1^{Ge'} \mathbf{X}_t^{Ge} + \theta_{00} + N * \theta_{01} + \theta'_1 \Phi \mathbf{X}_t + \bar{A}_N + \bar{\mathbf{B}}'_N \Phi^{Ge} \mathbf{X}_t^{Ge} \\ & + \bar{D}_N + \bar{\mathbf{E}}'_N \Phi \mathbf{X}_t + \frac{1}{2} [\lambda'_t \lambda_t - 2 \bar{\mathbf{B}}'_N \Sigma^{Ge} \lambda_t + \bar{\mathbf{B}}'_N \Sigma^{Ge} \Sigma^{Ge'} \bar{\mathbf{B}}_N \\ & - 2 \theta'_1 \Sigma \mathbf{J}' \lambda_t + 2 \theta'_1 \Sigma \mathbf{J}' \Sigma^{Ge'} \bar{\mathbf{B}}_N - 2 \bar{\mathbf{E}}'_N \Sigma \mathbf{J}' \lambda_t + 2 \bar{\mathbf{E}}'_N \Sigma \mathbf{J}' \Sigma^{Ge'} \bar{\mathbf{B}}_N \\ & + (\theta'_1 + \bar{\mathbf{E}}'_N) \Sigma \Sigma' (\theta'_1 + \bar{\mathbf{E}}'_N)']. \end{aligned}$$

Substituting equations (A-2) to (A-4) and rearranging the equation above, gives: $v_t^{N+1} = v_t^{N+1, Ge} + \bar{D}_{N+1} + \bar{\mathbf{E}}'_{N+1} \mathbf{X}_t$, where \bar{D}_{N+1} and $\bar{\mathbf{E}}'_{N+1}$ are given by equations (1-9) and (1-10). The proof for $N = 2$ is straightforward.

C Data: Quarterly to Monthly Frequency

Fiscal Balance, Debt and GDP monthly data are not available. Monthly data of these variables were constructed from their quarterly observations and from monthly data of some coincident indicators, such as energy consumption, unemployment rate, imports and exports.

Quarterly data consists of general government revenues, general government expenditures, net lending/borrowing and GDP. Monthly data consists of unemployment rate (UN), industrial production (IP), consumer price index (CPI), imports and exports (M and X), energy consumption in GWh (EN), and car registrations (CAR). Transformation method has two steps. In the first step, we run quarterly data regressions:

$$Z_t^q = \alpha + \beta_{IP}IP_t^q + \beta_{UN}UN_t^q + \beta_{CPI}CPI_t^q + \beta_M M_t^q + \beta_X X_t^q + \beta_{EN}EN_t^q + \beta_{CAR}CAR_t^q + \varphi_q + \varepsilon_t$$

where Z is general government net lending/borrowing, total revenues, total expenditure or GDP. The superscript q means quarterly observation data. In the second step, we make the following quarterly to monthly frequency transformation:

$$\hat{Z}_t^m = \hat{\alpha} + \hat{\beta}_{IP}IP_t^m + \hat{\beta}_{UN}UN_t^m + \hat{\beta}_{CPI}CPI_t^m + \hat{\beta}_M M_t^m + \hat{\beta}_X X_t^m + \hat{\beta}_{EN}EN_t^m + \hat{\beta}_{CAR}CAR_t^m + \hat{\varphi}_q,$$

where $\hat{\beta}$'s are estimates from quarterly data. \hat{Z}_t^m is the variable used in our estimations.

D Identification Hypothesis

We can rewrite equation (1-2) in the following structural version:

$$AX_t = \alpha_1 X_{t-1} + Bu_t \quad (A4.1)$$

where A , α_1 and B are $K \times K$ matrices, α_{00} is a K -dimensional vector and $\text{Var}(u_t) = I$. If A is an invertible matrix, (A4.1) can be rewritten as: $X_t = A^{-1}\alpha_1 X_{t-1} + A^{-1}Bu_t$, where, in an analogy to equation (1-5), $A^{-1}\alpha_1 = \Phi$ and $A^{-1}B = \Sigma$. Following Bonomo and Lowenkron (2008), it is assumed that Σ is a lower triangular Cholesky matrix such that $\text{Cov}(\Sigma u_t \Sigma') = \Sigma \Sigma'$. Furthermore, as a consequence of the way the principal components of yields were constructed, it is natural to impose that macroeconomic and those factors have non-correlated contemporaneous relation. Also, this kind of restriction is a standard practice in the term structure models (see Dai and Singleton [2000], Ang and Piazzesi [2003] and Joslin, Priebsch and Singleton [2010], and among others). In the estimation of the dynamic of the state equation for Germany, we impose the following restrictions on

$$A \text{ and } B: A^{Ge} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21}^{Ge} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{31}^{Ge} & a_{32}^{Ge} & 1 & 0 & 0 & 0 & 0 & 0 \\ a_{41}^{Ge} & a_{42}^{Ge} & a_{43}^{Ge} & 1 & 0 & 0 & 0 & 0 \\ a_{51}^{Ge} & a_{52}^{Ge} & a_{53}^{Ge} & a_{54}^{Ge} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{65}^{Ge} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{75}^{Ge} & a_{76}^{Ge} & 1 & 0 \\ 0 & 0 & 0 & 0 & a_{85}^{Ge} & a_{86}^{Ge} & a_{87}^{Ge} & 1 \end{pmatrix}_{8 \times 8};$$

$$B^{Ge} = \begin{pmatrix} b_{11}^{Ge} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{22}^{Ge} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33}^{Ge} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{44}^{Ge} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55}^{Ge} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{66}^{Ge} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_{77}^{Ge} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{88}^{Ge} \end{pmatrix}_{8 \times 8}$$

We impose the following restrictions on A and B in the estimation of country

i 's model:

$$A^i = \begin{pmatrix} A^{Ge} & 0_{8 \times 7} \\ Asp^i \end{pmatrix} \quad B^i = \begin{pmatrix} B^{Ge} & 0_{8 \times 7} \\ 0_{7 \times 8} & B^i \end{pmatrix}, \text{ where}$$

$$Asp^i = \begin{pmatrix} a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & 1 & 0 & 0 & 0 & 0 \\ a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & 1 & 0 & 0 & 0 \\ a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & a_{11}^i & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{11}^i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{11}^i & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{11}^i & a_{11}^i & 1 \end{pmatrix}_{7 \times 15},$$

and

$$B^i = \begin{pmatrix} b_{11}^i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{11}^i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{11}^i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{11}^i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{11}^i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{11}^i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_{11}^i \end{pmatrix}_{7 \times 7}.$$

The ordering of variables - with central economy coming first - follows Bonomo and Lowenkron (2008). However, in times of sovereign credit crisis, it is not clear whether the central economy's variables are more exogenous than the variables of the countries that are originating the crisis. Again, we expect this concern is reduced by the inclusion of the Moody's risk indicator in the model for Germany. The correlation between Moody's High Yield and yields on bonds of specially longer maturities is large: for 10-year bond yields, it is 0.83 for Italy, 0.79 for Spain and 0.61 for Greece (see scatter-plots in Figure 3).

It still remains to discuss our option regarding ordering of country-specific variables. It is usual to consider activity more exogenous than inflation: Joslin, Singleton and Priebish (2010) and Perotti (2004) are examples. We follow Favero (2002) and identify fiscal shocks by imposing that the activity and inflation indicators do not contemporaneously react to it. We do not identify non-policy shocks. The fact that our activity indicator is industrial production instead of GDP makes us comfortable about the ordering choice. As noted by Perotti (2005), ordering the fiscal policy instrument after GDP is questionable because government spending is a component of GDP. Therefore, this assumption would impose an implicit assumption of exactly 100 percent crowding out contemporaneously on private GDP. Finally, assuming that deficit is more exogenous than debt-GDP ratio seems reasonable given the definition of these variables.

It is important to stress that the over identification restrictions on matrices \mathbf{A} and \mathbf{Asp} (more specifically, the zero restrictions on the lower-left blocks of the matrices) are implications of method used to find the principal components of yields and spreads: they were estimated from an eigenvalue decomposition of the observed yields (spreads) covariance matrix, orthogonalized with respect to the macroeconomics and risk factors. Therefore, the contemporaneous correlation between macroeconomic variables and these principal components is zero.

E Bootstrap procedure

The non-parametric Bootstrap procedure follows from Horowitz (2001) and has two steps. In the first step, we estimate the model (the VAR equation (2) and the pricing equations (1-3) and (1-11)) and, for each equation, we save the estimated parameters $\widehat{\Omega}$ and the vector of residuals \mathbf{u}_t . The second step is composed of 400 loops. In the beginning of each loop, we draw a random sample with replacement from the vector residuals \mathbf{u}_t of the same size of \mathbf{u}_t , creating \mathbf{u}_t^* . Next, using the original regressors, the parameters estimated in the first step $\widehat{\Omega}$ and the residuals \mathbf{u}_t^* , we construct the new dependent variable vector. We then re-estimate the model using the same regressors and the new set of dependent variables. We save each of the 400 sets of estimated parameters and based on them, we calculate the mean and standard deviation of the parameters.

There is one difference in the resampling of residuals between the VAR equation (1-2) and the pricing equations (1-3) and (1-11). This bootstrap procedure requires that errors are not autocorrelated. We do not reject the hypothesis (with 95% level of significance) that the pricing errors are autocorrelated of order one in the equations of bonds of longer maturities (it was the case of the 10-year bond for Italy, the 5-, 7, and 10- year bond for Spain and the 2-, 3-, 4-, 5-, 7- and 10-year bond for Greece). For this reason, the bootstrap procedure on the pricing equations has an intermediate step. After the first step, we estimate an AR(1) equation for the pricing residuals \mathbf{u}_t , $\mathbf{u}_t = \alpha + \beta' \mathbf{u}_{t-1} + \xi_t$ and save the AR(1) parameters ($\widehat{\alpha}$ and $\widehat{\beta}$). Then, we draw a random sample with replacement from residuals ξ_t and create a new time series ξ_t^* . The new set of resampled residuals \mathbf{u}_t is $\mathbf{u}_t^* = \widehat{\alpha} + \widehat{\beta}' \mathbf{u}_{t-1} + \xi_t^*$. This way, if the vector of residual is autocorrelated of order one, our bootstrap procedure is not invalidated. The second step is unchanged.

F Ricatti difference equations

If $M_{t+1}V_{t+1}^N$ in equation (2-2) follows a lognormal distribution, the restriction of no arbitrage implies that

$$v_t^N = \log(V_t^N) = \bar{A}_N + \bar{B}'_N X_t^{US},$$

where

$$\bar{A}_{N+1} = \bar{A}_N - \bar{B}'_N \Sigma \lambda_0 + \frac{1}{2} \bar{B}'_N \Sigma \Sigma' \bar{B}_N - \delta_0^{US} \quad (\text{F-1})$$

$$\bar{B}'_{N+1} = B'_N (\Phi^{US} - \Sigma \lambda'_1) - \delta_1^{US'}, \quad (\text{F-2})$$

with $\bar{A}_1 = -\delta_0^{US}$ and $\bar{B}_1 = -\delta_1^{US}$. The yield y_t^N at t of a bond with maturity N is given by:

$$y_t^N = \frac{-\log(V_t^N)}{N} = A_N + B'_N X_t^{US}, \quad (\text{F-3})$$

where $A_N = \frac{-\bar{A}_N}{N}$ and $B_N = \frac{-\bar{B}_N}{N}$.

G Proof of equation (2-10)

The proof of equation (2-10) has two steps: In the first one, we prove that the short rate spread $s_t^{1,BRUS} \equiv y_t^{1,BRUS} - y_t^{1,US}$ is linear in X_t^{BR} ; in the second step, we show that, given that the short rate spread is linear in X_t^{BR} and the dynamics of the U.S. yield curve is given by equation (F-3), spreads of other maturities are given by equations (2-10) to (2-13).

First step: the short rate spread $s_t^{1,BRUS} \equiv y_t^{1,BRUS} - y_t^{1,US}$ is linear in X_t^{BR} . Consider $V_t^{N,BRUS}$ the price at t of a bond that pays no coupon and makes a payment of 1 USD at time $t+N$. For $N=0$, $V_t^{N,BRUS} = 1$. So equation (B-2) can be written as

$$V_t^{1,BRUS} = \mathbb{E}[M_{t+1}\Theta_{t+1}^0]. \quad (\text{G-1})$$

If $M_{t+1}\Theta_{t+1}^N V_{t+1}^N$ follows a lognormal distribution, equation (G-1) gives $v_t^{1,BRUS} = \mathbb{E}[m_{t+1} + \theta t + 1^0] + 1/2\text{Var}[m_{t+1} + \theta_{t+1}^0]$, where lowercase letters denote variables in log. Substituting equations (2-3) and (2-9), we have:

$$\begin{aligned} v_t^{1,BRUS} &= \mathbb{E}\left[-\frac{1}{2}\lambda_t'\lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US} - \lambda_t'\varepsilon_{t+1}^{US} + \theta_{00} + 0 * \theta_{01}\right. \\ &\quad \left.+ \theta_1'(\Phi^{BR} X_t^{BR} + \Sigma^{BR}\varepsilon_{t+1}^{BR})\right] + 1/2\text{Var}\left[-\frac{1}{2}\lambda_t'\lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US}\right. \\ &\quad \left.- \lambda_t'\varepsilon_{t+1}^{US} + \theta_{00} + 0 * \theta_{01} + \theta_1'(\Phi^{BR} X_t^{BR} + \Sigma^{BR}\varepsilon_{t+1}^{BR})\right] \\ &= \mathbb{E}\left[-\frac{1}{2}\lambda_t'\lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + \theta_1'\Phi^{BR} X_t^{BR}\right] \\ &\quad + 1/2\text{Var}\left[-\lambda_t'\varepsilon_{t+1}^{US} + \theta_1'\Sigma^{BR}\varepsilon_{t+1}^{BR}\right] \\ &= -\frac{1}{2}\lambda_t'\lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + \theta_1'\Phi^{BR} X_t^{BR} \\ &\quad + 1/2\text{Var}\left[(-\lambda_t'J + \theta_1'\Sigma^{BR})\varepsilon_{t+1}^{BR}\right], \end{aligned}$$

where J is a selection $(K^{US} \times K)$ -dimensional matrix such that $\varepsilon_t^{US} = J\varepsilon_t^{BR}$. Note that $JJ' = I_{K^{US}}$. Therefore, $v_t^{1,BRUS} = -\delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + \theta_1'\Phi X_t - \theta_1'\Sigma^{BR} J'\lambda_t + 0.5 * \theta_1'\Sigma^{BR}\Sigma^{BR'}\theta_1$. Substituting equation (2-4), $v_t^{1,BRUS} = -\delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + \theta_1'\Phi X_t^{BR} - \theta_1'\Sigma^{BR} J'(\lambda_0 + \lambda_1' X_t^{US}) + 0.5 * \theta_1'\Sigma^{BR}\Sigma^{BR'}\theta_1$. We know from equation (2-1) that $v_t^{1,US} = -r_t^{US} = -\delta_0^{US} - \delta_1^{US'} X_t^{US}$. So, $v_t^{1,BRUS} = v_t^{1,US} + [\theta_{00} - \theta_1'\Sigma^{BR} J'\lambda_0 + 0.5 * \theta_1'\Sigma^{BR}\Sigma^{BR'}\theta_1] + [\theta_1'(\Phi - \Sigma^{BR} J'\lambda_1' J)]X_t^{BR}$.

It means that:

$$v_t^{1,BRUS} = v_t^{1,US} + \bar{D}_1^{BRUS} + \bar{E}_1^{BRUS'} X_t^{BR}, \quad (G-2)$$

where $\bar{D}_1^{BRUS} = [[\theta_{00} - \theta_1' \Sigma^{BR} J' \lambda_0 + 0.5 * \theta_1' \Sigma^{BR} \Sigma^{BR'} \theta_1]]$, and $\bar{E}_1^{BRUS'} = [\theta_1' (\Phi - \Sigma^{BR} J' \lambda_1' J)]$. Equation (G-2) can be written the following way: $s_t^{1,BRUS} = y_t^{1,BRUS} - y_t^{1,US} = -\frac{v_t^{1,BRUS}}{1} + \frac{v_t^{1,US}}{1} = D_1^{BRUS} + E_1^{BRUS'} X_t^{BRUS}$, where $D_1^{BRUS} = -\frac{\bar{D}_1^{BRUS}}{1}$ and $E_1^{BRUS} = -\frac{\bar{E}_1^{BRUS'}}{1}$.

Second step: given that the short rate spread is linear in X_t^{BR} and the dynamics of U.S. yield curve is given by equation (F-3), spreads of other maturities are given by equations (2-10) to (2-13). From equation (B-2), and assuming that $M_{t+1} \Theta_{t+1}^N V_{t+1}^{N,BRUS}$ follows a lognormal distribution, we have

$$v_t^{N+1,BRUS} = \mathbb{E}[m_{t+1} + \theta_{t+1}^N + v_{t+1}^{N,BRUS}] + 1/2 \text{Var}[m_{t+1} + \theta_{t+1}^N + v_{t+1}^{N,BRUS}]. \quad (G-3)$$

Take $N = 1$. Substituting equations (G-2), (2-3), (2-4), (2-7), and (2-9) into equation (G-3) we have:

$$\begin{aligned} v_t^{N+1,BRUS} = & -\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + N * \theta_{01} + \theta_1' \Phi^{BR} X_t^{BR} + \bar{A}_N \\ & + \bar{B}_N' \Phi^{US} X_t^{US} + \bar{D}_N + \bar{E}_N' X_t^{BR} \\ & + 1/2 \text{Var}[(\bar{B}_N' \Sigma^{US} - \lambda_t') J + (\theta_1 + \bar{E}_N^{BRUS})' \Sigma^{BR}] \varepsilon_{t+1}^{BR}. \end{aligned} \quad (G-4)$$

Opening only $\text{Var}[(\bar{B}_N' \Sigma^{US} - \lambda_t') J + (\theta_1 + \bar{E}_N^{BRUS})' \Sigma^{BR}] \varepsilon_{t+1}^{BR}$, we have

$$\begin{aligned} \text{Var}[(\bar{B}_N' \Sigma^{US} - \lambda_t') J + (\theta_1 + \bar{E}_N^{BRUS})' \Sigma^{BR}] \varepsilon_{t+1}^{BR} = & \lambda_t' \lambda_t - 2 \bar{B}_N' \Sigma^{US} \Sigma^{US'} \bar{B}_N \\ & - 2 \theta_1' \Sigma^{BR} J' \lambda_t \\ & + 2 \theta_1' \Sigma^{BR} J' \Sigma^{US'} \bar{B}_N - 2 \bar{E}_N' \Sigma^{BR} J' \lambda_t \\ & + 2 \bar{E}_N' \Sigma^{BR} J' \Sigma^{US'} \bar{B}_N + (\theta_1 + \bar{E}_N)' \Sigma^{BR} \Sigma^{BR'} (\theta_1 + \bar{E}_N). \end{aligned}$$

Substituting into equation (G-4),

$$\begin{aligned} v_t^{N+1,BRUS} = & -\frac{1}{2} \lambda_t' \lambda_t - \delta_0^{US} - \delta_1^{US'} X_t^{US} + \theta_{00} + N * \theta_{01} + \theta_1' \Phi^{BR} X_t^{BR} \\ & + \bar{A}_N + \bar{B}_N' \Phi^{US} X_t^{US} + \bar{D}_N + \bar{E}_N' X_t^{BR} \\ & + \frac{1}{2} [\lambda_t' \lambda_t - 2 \bar{B}_N' \Sigma^{US} \Sigma^{US'} \bar{B}_N - 2 \theta_1' \Sigma^{BR} J' \lambda_t + 2 \theta_1' \Sigma^{BR} J' \Sigma^{US'} \bar{B}_N \\ & - 2 \bar{E}_N' \Sigma^{BR} J' \lambda_t + 2 \bar{E}_N' \Sigma^{BR} J' \Sigma^{US'} \bar{B}_N \\ & + (\theta_1 + \bar{E}_N)' \Sigma^{BR} \Sigma^{BR'} (\theta_1 + \bar{E}_N)]. \end{aligned}$$

Substituting equations (2-9) and (F-1) to (F-3), and rearranging the equation above, gives: $v_t^{N+1,BRUS} = v_t^{N+1,US} + \bar{D}_{N+1}^{BRUS} + \bar{E}_{N+1}^{BRUS'} X_t^{BR}$, where \bar{D}_{N+1}^{BRUS} and $\bar{E}_{N+1}^{BRUS'}$ are given by equations (2-11) and (2-13). The proof for $N \geq 2$ is straightforward.

H Proof of equation (2-16)

As π_t , as well as θ_t , is linear in X_t^{BR} , the proof of equation (2-16) is analog to the proof of equation (2-10). Recall that

$$v_t^{N+1,BRUS} = \mathbb{E}[m_{t+1} + \theta_{t+1}^N + v_{t+1,BRUS}^N] + \frac{1}{2}\mathbb{V}ar[m_{t+1} + \theta_{t+1}^N + v_{t+1,BRUS}^N]$$

and

$$v_t^{N+1,BRRS} = \mathbb{E}[m_{t+1} + \gamma_{t+1}^N + v_{t+1,BRRS}^N] + \frac{1}{2}\mathbb{V}ar[m_{t+1} + \gamma_{t+1}^N + v_{t+1,BRRS}^N]$$

where $\gamma_{t+1} = \theta_{t+1}^N + \pi_{t+1}$. We can write

$$\gamma_{t+1}^N = \gamma_{00} + \gamma_{01} * N + \gamma_1 X_{t+1}^{BR},$$

where $\gamma_{00} = \theta_{00} + \pi_0$, $\gamma_{01} = \theta_{01}$, and $\gamma_1 = \theta_1 + \pi_1$. Therefore, if we substitute θ for γ in the proof of equation (2-10), we will prove equation (2-16).

I Proofs of Chapter 3

Proof of Theorem 1. Write the event

$$\begin{aligned}
\{\widehat{M} \neq M^*\} &\Leftrightarrow \{\widehat{M} \neq M^* \cap M > M^*\} \cup \{\widehat{M} \neq M^* \cap M < M^*\} \\
&\Rightarrow \{\widehat{M} > M^* \cap M > M^*\} \cup \{\widehat{M} < M^* \cap M < M^*\} \\
&\Leftrightarrow \{\text{IC}(\widehat{M}) < \text{IC}(M^*) \cap M > M^*\} \cup \{\text{IC}(\widehat{M}) < \text{IC}(M^*) \cap M < M^*\} \\
&\Rightarrow \left[\bigcup_{M=M^*+1}^{\overline{M}} \{\text{IC}(M) < \text{IC}(M^*)\} \right] \cup \left[\bigcup_{M=1}^{M^*-1} \{\text{IC}(M) < \text{IC}(M^*)\} \right] \\
&=: A \cup B.
\end{aligned}$$

We show that $\mathbb{P}(A \cup B) \rightarrow 0$ as $N \rightarrow \infty$. It is clear that

$$\{\text{IC}(M) < \text{IC}(M^*)\} \Leftrightarrow \{\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M) > \lambda_N(M) - \lambda_N(M^*)\}.$$

Applying the Markov inequality to the right hand side of the above equation, we have, for $M > M^*$,

$$P\left(\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M) > \lambda_N(M) - \lambda_N(M^*)\right) \leq \frac{\mathbb{E}\left|\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M)\right|}{\lambda_N(M) - \lambda_N(M^*)}.$$

Using the triangular inequality, we obtain

$$\begin{aligned}
\mathbb{E}\left|\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M)\right| &\leq \max_{\widehat{\psi} \in \Psi_{N, M^*}} \min_{\psi \in \Psi_M^*} \mathbb{E}\left|\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\psi, M^*)\right| \\
&\quad + \max_{\widehat{\psi} \in \Psi_{N, M}} \min_{\psi \in \Psi_M^*} \mathbb{E}\left|\mathcal{Q}(\widehat{\psi}, M) - \mathcal{Q}(\psi, M)\right| \\
&\quad + \min_{\psi \in \Psi_M^*} \mathbb{E}\left|\mathcal{Q}(\psi, M) - \mathcal{Q}(\psi, M^*)\right| = A_1 + A_2 + A_3.
\end{aligned}$$

Since $M > M^*$, Assumptions 5(1) and 5(5) guarantee that $A_1 + A_2 \leq 2ck_{\psi, M}^2$. $A_3 = 0$ as, by definition, $\mathbb{E}[\mathcal{Q}(\cdot, M^*)]$ is minimum over all values of M .

Using the union bound on A we have

$$P(A) \leq 2c \sum_{M=M^*+1}^{\overline{M}} \frac{k_{\psi, M}^2}{\lambda_N(M) - \lambda_N(M^*)} \rightarrow 0$$

as $N \rightarrow \infty$.

Now assume that $M < M^*$ and write for $\psi \in \Psi_M^*$

$$\begin{aligned} \frac{1}{N} \left[\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M) \right] &= \frac{1}{N} \left\{ \mathcal{Q}(\widehat{\psi}, M^*) - \mathbb{E} \left[\mathcal{Q}(\widehat{\psi}, M^*) \right] \right\} \\ &\quad - \frac{1}{N} \left\{ \mathcal{Q}(\widehat{\psi}, M) - \mathbb{E} \left[\mathcal{Q}(\widehat{\psi}, M) \right] \right\} \\ &\quad + \frac{1}{N} \mathbb{E} \left[\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\psi, M^*) \right] \\ &\quad - \frac{1}{N} \mathbb{E} \left[\mathcal{Q}(\widehat{\psi}, M) - \mathcal{Q}(\psi, M) \right] \\ &\quad - \frac{1}{N} \mathbb{E} \left[\mathcal{Q}(\psi, M) - \mathcal{Q}(\psi, M^*) \right] \\ &= o_p(1) - K_M, \end{aligned}$$

where $K_M = \frac{1}{N} \mathbb{E} [\mathcal{Q}(\psi, M) - \mathcal{Q}(\psi, M^*)]$.

The first line of the above expression is clearly $o_p(1)$ by the law of large numbers, the second line is $o(1)$ by Assumption 5(5), and $K_M > 0$ by the definition of M^* . From Assumption 5(3) $\frac{1}{N} [\lambda_N(M^*) - \lambda_N(M)] \rightarrow 0$ as $N \rightarrow \infty$. Therefore, the set

$$\{\mathcal{Q}(\widehat{\psi}, M^*) - \mathcal{Q}(\widehat{\psi}, M) > \lambda_N(M) - \lambda_N(M^*)\} \rightarrow \emptyset.$$

Since $M^* < \infty$ by definition, it follows from the union bound that $P(B) \rightarrow 0$ and the theorem is proved. \square

J Countries in the sample

Countries, years and the number of industries observed for each country-year pair are shown in Tables J.1 to J.4. To be present in the sample, it is necessary that data on productivity is available in t and in $t - 10$. But recall that the use of variables not from UNIDO's INDSTAT 4 data base reduces the number of countries in our data set because these variables are missing for some countries.

Trade openness reduces the number of countries to 48, and the number of observations to 13,265 (Table J.1); the executive constraints indicator reduces the number of countries to 38, and the number of observations to 11,363 (Table J.2); the years of schooling indicator reduces the number of countries to 43, and the number of observations to 12,499 (Table J.3). Finally, the use of all the three indicators reduces the number of countries to 37, and the number of observations to 11,098 (Table J.4).

Table J.1: Observed Countries, years and number of industries. Trade openness and INDSTST4 available data in common.

Australia	Years	-	-	-	2003	-	-	-	-	All
	No. of Industries	-	-	-	9	-	-	-	-	9
Austria	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	59	64	62	62	62	98	10	92	509
Belgium	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	99	97	103	299
Brazil	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	35	35	70
Bulgaria	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	27	28	55
Canada	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	106	106	106	106	102	100	102	101	829
Czech Republic	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	58	65	55	178
Denmark	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	73	72	73	218
Ecuador	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	95	95	91	281
Ethiopia	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	10	40	39	40	41	40	44	43	297
Eritrea	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	34	34	34	30	29	34	195
Finland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	11	11	11	10	107	106	109	365
France	Years	2000	2001	2002	2003	2004	-	2006	2007	All
	No. of Industries	104	104	100	98	104	-	120	122	752
Greece	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	31	32	29	28	30	150
Hungary	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	76	78	78	78	310
Iran	Years	-	-	-	-	2004	2005	-	-	All
	No. of Industries	-	-	-	-	121	119	-	-	240
Ireland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	60	47	62	57	54	47	37	364
Israel	Years	-	-	-	-	-	2005	2006	-	All
	No. of Industries	-	-	-	-	-	24	24	-	48
Italy	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	123	121	121	121	119	120	725
Japan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	119	120	119	119	477
Jordan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	74	74	76	74	298
Republic of Korea	Years	2000	2001	2002	2003	2004	2005	2006	-	All
	No. of Industries	117	127	126	126	126	126	126	-	874
Latvia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	83	80	80	83	45	371
Luxembourg	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	4	4	4	5	5	18	16	17	73
China, Macao SAR	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	3	3
Malta	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	56	58	59	173
Mauritius	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	7	7
Mongolia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	3	4	7	6	7	27
Oman	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	50	53	55	53	54	265
Netherlands	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	75	84	81	240
New Zealand	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	4	4	8
Norway	Years	-	2001	2002	2003	2004	2005	2006	-	All
	No. of Industries	-	81	70	69	105	102	96	-	523
Peru	Years	-	-	-	-	-	2005	2006	-	All
	No. of Industries	-	-	-	-	-	9	9	-	18
Philippines	Years	-	-	-	-	-	-	2006	-	All
	No. of Industries	-	-	-	-	-	-	21	-	21
Poland	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	11	12	12	12	110	112	269
Portugal	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	66	62	128
Romania	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	9	9	9	8	8	7	6	7	63
Singapore	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	71	71	70	53	59	58	61	443
Slovakia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	74	82	61	68	73	358
Slovenia	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	82	78	79	239
South Africa	Years	-	2001	-	2003	-	-	2006	-	All
	No. of Industries	-	9	-	9	-	-	11	-	29
Spain	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	123	123	123	123	123	615
Sweden	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	95	92	90	93	96	85	83	89	723
Thailand	Years	-	-	-	-	-	-	2006	-	All
	No. of Industries	-	-	-	-	-	-	121	-	121
Turkey	Years	-	-	2002	2003	2004	2005	2006	-	All
	No. of Industries	-	-	11	99	99	98	11	-	318
The f. Yugosl. Rep of Macedonia	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	6	6
United Kingdom	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	109	113	113	117	116	568
United States of America	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	113	113
All Countries	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	504	778	914	1507	1912	2487	2701	2462	13265

Table J.2: Observed Countries, years and number of industries. Executive constraint and INDSTST4 available data in common.

Australia	Years	-	-	-	2003	-	-	-	-	All
	No. of Industries	-	-	-	9	-	-	-	-	9
Austria	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	59	64	62	62	62	98	10	92	509
Belgium	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	99	97	103	299
Brazil	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	35	35	70
Bulgaria	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	27	28	55
Canada	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	106	106	106	106	102	100	102	101	829
Czech Republic	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	58	65	55	178
Denmark	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	73	72	73	218
Ecuador	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	95	95	91	281
Ethiopia	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	10	40	39	40	41	40	44	43	297
Eritrea	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	34	34	34	30	29	34	195
Finland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	11	11	11	10	107	106	109	365
France	Years	2000	2001	2002	2003	2004	-	2006	2007	All
	No. of Industries	104	104	100	98	104	-	120	122	752
Greece	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	31	32	29	28	30	150
Hungary	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	76	78	78	78	310
Iran	Years	-	-	-	-	2004	2005	-	-	All
	No. of Industries	-	-	-	-	121	119	-	-	240
Ireland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	60	47	62	57	54	47	37	364
Israel	Years	-	-	-	-	-	2005	2006	-	All
	No. of Industries	-	-	-	-	-	24	24	-	48
Italy	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	123	121	121	121	119	120	725
Japan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	119	120	119	119	477
Jordan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. of Industries	-	-	-	-	74	74	76	74	298
Republic of Korea	Years	2000	2001	2002	2003	2004	2005	2006	-	All
	No. of Industries	117	127	126	126	126	126	126	-	874
Latvia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	83	80	80	83	45	371
Luxembourg	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	4	4	4	5	5	18	16	17	73
China, Macao SAR	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	3	3
Malta	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	56	58	59	173
Mauritius	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	7	7
Mongolia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	3	4	7	6	7	27
Oman	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	50	53	55	53	54	265
Netherlands	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	75	84	81	240
New Zealand	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	4	4	8
Norway	Years	-	2001	2002	2003	2004	2005	2006	-	All
	No. of Industries	-	81	70	69	105	102	96	-	523
Peru	Years	-	-	-	-	-	2005	2006	-	All
	No. of Industries	-	-	-	-	-	9	9	-	18
Philippines	Years	-	-	-	-	-	-	2006	-	All
	No. of Industries	-	-	-	-	-	-	21	-	21
Poland	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	11	12	12	12	110	112	269
Portugal	Years	-	-	-	-	-	-	2006	2007	All
	No. of Industries	-	-	-	-	-	-	66	62	128
Romania	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	9	9	9	8	8	7	6	7	63
Singapore	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	-	71	71	70	53	59	58	61	443
Slovakia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	74	82	61	68	73	358
Slovenia	Years	-	-	-	-	-	2005	2006	2007	All
	No. of Industries	-	-	-	-	-	82	78	79	239
South Africa	Years	-	2001	-	2003	-	-	2006	-	All
	No. of Industries	-	9	-	9	-	-	11	-	29
Spain	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	123	123	123	123	123	615
Sweden	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	95	92	90	93	96	85	83	89	723
Thailand	Years	-	-	-	-	-	-	2006	-	All
	No. of Industries	-	-	-	-	-	-	121	-	121
Turkey	Years	-	-	2002	2003	2004	2005	2006	-	All
	No. of Industries	-	-	11	99	99	98	11	-	318
The f. Yugosl. Rep of Macedonia	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	6	6
United Kingdom	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. of Industries	-	-	-	109	113	113	117	116	568
United States of America	Years	-	-	-	-	-	-	-	2007	All
	No. of Industries	-	-	-	-	-	-	-	113	113
All Countries	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	504	778	914	1507	1912	2487	2701	2462	13265

Table J.3: Observed Countries, years and number of industries. Years of schooling and INDSTST4 available data in common.

Australia	Years	-	-	-	2003	-	-	-	-	All
	No. Of Industries	-	-	-	9	-	-	-	-	9
Austria	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	59	64	62	62	62	98	10	92	509
Belgium	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	99	97	103	299
Brazil	Years	-	-	-	-	-	-	2006	2007	All
	No. Of Industries	-	-	-	-	-	-	35	35	70
Bulgaria	Years	-	-	-	-	-	-	2006	2007	All
	No. Of Industries	-	-	-	-	-	-	27	28	55
Canada	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	106	106	106	106	102	100	102	101	829
Czech Republic	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	58	65	55	178
Denmark	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	73	72	73	218
Ecuador	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	95	95	91	281
Finland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	-	11	11	11	10	107	106	109	365
France	Years	2000	2001	2002	2003	2004	-	2006	2007	All
	No. Of Industries	104	104	100	98	104	-	120	122	752
Greece	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	31	32	29	28	30	150
Hungary	Years	-	-	-	-	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	-	76	78	78	78	310
Iran	Years	-	-	-	-	2004	2005	-	-	All
	No. Of Industries	-	-	-	-	121	119	-	-	240
Ireland	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	-	60	47	62	57	54	47	37	364
Israel	Years	-	-	-	-	-	2005	2006	-	All
	No. Of Industries	-	-	-	-	-	24	24	-	48
Italy	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	123	121	121	121	119	120	725
Japan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	-	119	120	119	119	477
Jordan	Years	-	-	-	-	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	-	74	74	76	74	298
Republic of Korea	Years	2000	2001	2002	2003	2004	2005	2006	-	All
	No. Of Industries	117	127	126	126	126	126	-	-	874
Latvia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	83	80	80	83	45	371
Luxembourg	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	4	4	4	5	5	18	16	17	73
Malta	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	56	58	59	173
Mauritius	Years	-	-	-	-	-	-	-	2007	All
	No. Of Industries	-	-	-	-	-	-	-	7	7
Mongolia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	3	4	7	6	7	27
Netherlands	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	75	84	81	240
New Zealand	Years	-	-	-	-	-	-	2006	2007	All
	No. Of Industries	-	-	-	-	-	-	4	4	8
Norway	Years	-	2001	2002	2003	2004	2005	2006	-	All
	No. Of Industries	-	81	70	69	105	102	96	-	523
Peru	Years	-	-	-	-	-	2005	2006	-	All
	No. Of Industries	-	-	-	-	-	9	9	-	18
Philippines	Years	-	-	-	-	-	-	2006	-	All
	No. Of Industries	-	-	-	-	-	-	21	-	21
Poland	Years	-	-	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	11	12	12	12	110	112	269
Portugal	Years	-	-	-	-	-	-	2006	2007	All
	No. Of Industries	-	-	-	-	-	-	66	62	128
Romania	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	9	9	9	8	8	7	6	7	63
Singapore	Years	-	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	-	71	71	70	53	59	58	61	443
Slovakia	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	74	82	61	68	73	358
Slovenia	Years	-	-	-	-	-	2005	2006	2007	All
	No. Of Industries	-	-	-	-	-	82	78	79	239
South Africa	Years	-	2001	-	2003	-	-	2006	-	All
	No. Of Industries	-	9	-	9	-	-	11	-	29
Spain	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	123	123	123	123	123	615
Sweden	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. Of Industries	95	92	90	93	96	85	83	89	723
Thailand	Years	-	-	-	-	-	-	2006	-	All
	No. Of Industries	-	-	-	-	-	-	121	-	121
Turkey	Years	-	-	2002	2003	2004	2005	2006	-	All
	No. Of Industries	-	-	11	99	99	98	11	-	318
United Kingdom	Years	-	-	-	2003	2004	2005	2006	2007	All
	No. Of Industries	-	-	-	109	113	113	117	116	568
United States of America	Years	-	-	-	-	-	-	-	2007	All
	No. Of Industries	-	-	-	-	-	-	-	113	113
All Countries	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
	No. of Industries	494	738	841	1383	1784	2362	2575	2322	12499

Table J.4: Observed Countries, years and number of industries. Trade openness, executive constraint, years of schooling and INDSTST4 available data in common.

	Years	-	-	-	2003	-	-	-	-	All
Australia	No. Of Industries	-	-	-	9	-	-	-	-	9
	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
Austria	No. Of Industries	59	64	62	62	62	98	10	92	509
	Years	-	-	-	-	-	2005	2006	2007	All
Belgium	No. Of Industries	-	-	-	-	-	99	97	103	299
	Years	-	-	-	-	-	-	2006	2007	All
Brazil	No. Of Industries	-	-	-	-	-	-	35	35	70
	Years	-	-	-	-	-	-	2006	2007	All
Bulgaria	No. Of Industries	-	-	-	-	-	-	27	28	55
	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
Canada	No. Of Industries	106	106	106	106	102	100	102	101	829
	Years	-	-	-	-	-	2005	2006	2007	All
Denmark	No. Of Industries	-	-	-	-	-	73	72	73	218
	Years	-	-	-	-	-	2005	2006	2007	All
Ecuador	No. Of Industries	-	-	-	-	-	95	95	91	281
	Years	-	2001	2002	2003	2004	2005	2006	2007	All
Finland	No. Of Industries	-	11	11	11	10	107	106	109	365
	Years	2000	2001	2002	2003	2004	-	2006	2007	All
France	No. Of Industries	104	104	100	98	104	-	120	122	752
	Years	-	-	-	2003	2004	2005	2006	2007	All
Greece	No. Of Industries	-	-	-	31	32	29	28	30	150
	Years	-	-	-	-	2004	2005	2006	2007	All
Hungary	No. Of Industries	-	-	-	-	76	78	78	78	310
	Years	-	-	-	-	2004	2005	-	-	All
Iran	No. Of Industries	-	-	-	-	121	119	-	-	240
	Years	-	2001	2002	2003	2004	2005	2006	2007	All
Ireland	No. Of Industries	-	60	47	62	57	54	47	37	364
	Years	-	-	-	-	-	2005	2006	-	All
Israel	No. Of Industries	-	-	-	-	-	24	24	-	48
	Years	-	-	2002	2003	2004	2005	2006	2007	All
Italy	No. Of Industries	-	-	123	121	121	121	119	120	725
	Years	-	-	-	-	2004	2005	2006	2007	All
Japan	No. Of Industries	-	-	-	-	119	120	119	119	477
	Years	-	-	-	-	2004	2005	2006	2007	All
Jordan	No. Of Industries	-	-	-	-	74	74	76	74	298
	Years	2000	2001	2002	2003	2004	2005	2006	-	All
Republic of Korea	No. Of Industries	117	127	126	126	126	126	126	-	874
	Years	-	-	-	-	-	-	-	2007	All
Mauritius	No. Of Industries	-	-	-	-	-	-	-	7	7
	Years	-	-	-	2003	2004	2005	2006	2007	All
Mongolia	No. Of Industries	-	-	-	3	4	7	6	7	27
	Years	-	-	-	-	-	2005	2006	2007	All
Netherlands	No. Of Industries	-	-	-	-	-	75	84	81	240
	Years	-	-	-	-	-	-	2006	2007	All
New Zealand	No. Of Industries	-	-	-	-	-	-	4	4	8
	Years	-	2001	2002	2003	2004	2005	2006	-	All
Norway	No. Of Industries	-	81	70	69	105	102	96	-	523
	Years	-	-	-	-	-	2005	2006	-	All
Peru	No. Of Industries	-	-	-	-	-	9	9	-	18
	Years	-	-	-	-	-	-	2006	-	All
Philippines	No. Of Industries	-	-	-	-	-	-	21	-	21
	Years	-	-	2002	2003	2004	2005	2006	2007	All
Poland	No. Of Industries	-	-	11	12	12	12	110	112	269
	Years	-	-	-	-	-	-	2006	2007	All
Portugal	No. Of Industries	-	-	-	-	-	-	66	62	128
	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
Romania	No. Of Industries	9	9	9	8	8	7	6	7	63
	Years	-	2001	2002	2003	2004	2005	2006	2007	All
Singapore	No. Of Industries	-	71	71	70	53	59	58	61	443
	Years	-	2001	-	-	-	-	2006	-	All
South Africa	No. Of Industries	-	9	-	-	-	-	11	-	20
	Years	-	-	-	2003	2004	2005	2006	2007	All
Spain	No. Of Industries	-	-	-	123	123	123	123	123	615
	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
Sweden	No. Of Industries	95	92	90	93	96	85	83	89	723
	Years	-	-	-	-	-	-	2006	-	All
Thailand	No. Of Industries	-	-	-	-	-	-	121	-	121
	Years	-	-	2002	2003	2004	2005	2006	-	All
Turkey	No. Of Industries	-	-	11	99	99	98	11	-	318
	Years	-	-	-	2003	2004	2005	2006	2007	All
United Kingdom	No. Of Industries	-	-	-	109	113	113	117	116	568
	Years	-	-	-	-	-	-	-	2007	All
United States of America	No. Of Industries	-	-	-	-	-	-	-	113	113
	Years	2000	2001	2002	2003	2004	2005	2006	2007	All
All Countries	No. of Industries	490	734	837	1212	1617	2007	2207	1994	11098

K Estimation of equation (4-10)

To estimate equation 4-10, we create the following matrix of regressors $y'_{ij,t}$

$$y'_{ij,t} = \begin{bmatrix} y_{11,1990} & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & 0 & 0 & 0 & 0 & 0 & 0 \\ y_{I_1,19901,1990} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & y_{11,1997} & 0 & 0 & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & y_{I_1,19971,1997} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y_{1J,1990} & 0 & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & y_{I_J,1990J,1990} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & y_{1J,1997} \\ 0 & 0 & 0 & 0 & 0 & 0 & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & y_{I_J,1997J,1997} \end{bmatrix} .$$

In each column are listed the observations of all industries of industry $i = 1, \dots, I_{j,t}$, country $j = 1, \dots, J$, year $t = 1990, \dots, 1997$, where $I_{j,t}$ is the number of observed industries in country j , year t . We then regress the equation

$$\Delta y_{ijt} = \beta y'_{ij,t} + D_i + D_t + \varepsilon_{ijt},$$

and obtain one different β_{jt} for each country j , year t .

L Versions of tables 4.4 and 4.3 in exercises where the convergence coefficient is estimated in an equation with no dummies and with only period dummies.

Table (L.1) shows the results of the linear regression of $\hat{\beta}_{jt}$'s on various indicators, measured in decadal initial level. The coefficients $\hat{\beta}_{jt}$'s are estimated according to equation (4-11). It is the analog of table 4.3 for equations with no dummies (panel A) and with only period dummies (panel B).

Table L.2 presents the country-and-year mean partial effects of latitude, longitude, trade openness, years of schooling and executive constraints on the relative productivity growth. These values are the partial effects of one standard deviation increase in these variables on the estimated convergence coefficient. It is the analog of table 4.4 for equations with no dummies (panel A) and with only period dummies (panel B).

Table L.1: $\hat{\beta}_{jt}$'s from the linear model and selected variables in the first year of the decade (t_0). Estimated coefficients are multiplied by 100.

Panel A - Dependent Variable: $\hat{\beta}_{jt}$'s from the equation with no dummies.				
	(1)	(2)	(3)	(4)
Latitude	0.393 (8.99)***	0.113 (2.77)***	0.028 (0.71)	-0.032 (0.76)
Longitude	-0.036 (0.92)	0.017 (0.60)	-0.012 (0.48)	0.008 (0.35)
Openness, t_0	0.054 (2.27)**			-0.006 (0.27)
Executive Constraint, t_0		0.144 (3.52)***		0.086 (1.60)
Years of Schooling, t_0			0.241 (6.30)***	0.277 (6.65)***
Constant	-3.515 (83.98)***	-3.302 (100.17)***	-3.332 (103.37)***	-3.337 (82.53)***
Observations	152	121	135	116
R-squared	0.37	0.29	0.36	0.50
Panel B - Dependent Variable: $\hat{\beta}_{jt}$'s from the equation with period dummies.				
	(5)	(6)	(7)	(8)
Latitude	0.415 (4.65)***	0.206 (1.82)*	0.071 (0.62)	0.049 (0.36)
Longitude	-0.013 (0.17)	0.037 (0.46)	0.013 (0.17)	0.011 (0.14)
Openness, t_0	0.039 (0.80)			-0.012 (0.18)
Executive Constraint, t_0		0.138 (1.20)		-0.043 (0.24)
Years of Schooling, t_0			0.240 (2.17)**	0.352 (2.58)**
Constant	-3.459 (40.52)***	-3.263 (35.51)***	-3.292 (35.45)***	-3.246 (24.51)***
Observations	152	121	135	116
R-squared	0.14	0.08	0.08	0.12

Absolute value of t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table L.2: Country-and-year mean partial effects
 . Industry productivity growth regressions: 10-year growth rates.
 Dependent variable: growth rate of labor productivity over relevant period.
 Estimated coefficients and standard values are multiplied by 100.

Panel A - Equation with no dummies				
	(1)	(2)	(3)	(4)
Latitude	0.237 (0.06)***	0.066 (0.19)	-0.018 (0.09)	0.035 (0.13)
Longitude	-0.113 (0.05)**	0.003 (0.28)	0.008 (0.09)	-0.023 (0.07)
Openness	0.086 (0.04)**			0.046 (0.26)
Executive Constraint		0.072 (0.24)		0.056 (0.37)
Years of Schooling			0.255 (0.10)**	0.240 (0.11)**
Number of Sieves	5	4	3	5
Number of Observations	13,265	11,363	12,499	11,098
R-squared (underlying regression)	0.24	0.16	0.30	0.26
Panel B - Equation with period dummies				
	(5)	(6)	(7)	(8)
Latitude	0.262 (0.15)	0.085 (0.04)**	-0.015 (0.16)	-0.035 (0.10)
Longitude	-0.071 (0.25)	0.011 (0.03)	0.013 (0.03)	0.027 (0.06)
Openness	0.057 (0.18)			-0.003 (0.10)
Executive Constraint		0.052 (0.02)***		0.163 (0.15)
Years of Schooling			0.263 (0.05)***	0.225 (0.11)**
Number of Sieves	5	4	3	5
Number of Observations	13,265	11,363	12,499	11,098
R-squared (underlying regression)	0.29	0.22	0.34	0.30

Standard deviations in parenthesis based on Monte Carlo simulations (1000 repetitions)