Euro corporate bonds risk factors

Castagnetti, Carolina and Rossi, Eduardo

università di Pavia, università di Pavia

16 October 2008

Euro Corporate Bonds Risk Factors*

Carolina Castagnetti † Università di Pavia

Eduardo Rossi‡ Università di Pavia

October 16, 2008

Abstract

This paper investigates the determinants of credit spread changes in Euro-denominated bonds. Because credit spread changes can be easily viewed as an excess return on corporate bonds over treasury bonds, we adopt a factor model framework, inspired by the credit risk structural approach. We try to assess the relative importance of market and idiosyncratic factors in explaining the movements in credit spreads. We adopt a heterogeneous panel with a multifactor error model and propose a two-step estimation procedure which yields consistent estimates of unobserved factors. The analysis is carried out with a panel of monthly redemption yields on a set of corporate bonds for a time span of three years. Our results suggest that the Euro corporate market is driven by observable and unobservable factors. Where the latter are identified through a consistent estimation of individual and common observable effects. We observe that the factors predicted by the structural model are not as relevant as in the case of the US market. The empirical results also suggest that an unobserved common factor has a significant role in explaining the systematic changes in credit spreads. However, contrary to the American evidence, it cannot be identified as a market factor.

Keywords: Euro Corporate Bonds, Cross Section Dependence, Common Correlated Effects, Yield Curve.
JEL - Classification: G10,C33

---

*This is a revised version of a paper presented at the Panel Data Conference 2006, Robinson College, Cambridge UK, July 8-9 and at the ESEM, Vienna, August 2006. We are indebted to Paolo Baldessari and Cristian Grigatti for helpful comments. Most of the database used in this paper was collected when Carolina Castagnetti was working at Fideuram Investimenti SGR. We are grateful to the quantitative group and the risk management group of Fideuram Investimenti SGR for very useful discussions and suggestions. We are also grateful to Giovanni Urga for very helpful comments. We thank Phil Galdi for providing us with the Merryll Lynch Database. Financial support from PRIN 2004 is gratefully acknowledged.

†Dipartimento di Economia Politica e Metodi Quantitativi, Via San Felice 5, 27100 Pavia, Italy, `castca@eco.unipv.it, tel.++390382986217, corresponding author`

‡Dipartimento di Economia Politica e Metodi Quantitativi, `erossi@eco.unipv.it`

1 Introduction

The credit risk, or risk of default, of a bond arises for two reasons: both the magnitude and the timing of payoffs to investors may be uncertain. In other words, the risk of default of the issuer is accompanied by the recovery rate uncertainty. The effects of default risks on prices depend on how the default event is defined and the specification of the recovery in the event of a default.

Because of this uncertainty, corporate bonds should offer higher yields than comparable default-free bonds, i.e. government bonds. Consequently, a corporate bond trades at a lower price than a corresponding (in terms of maturity and coupon) government bond. The difference between the yield on the risky bond and the yield on the corresponding default-free bond is called the credit spread.

Theoretical credit risk models tackle the default risk in different ways. Structural models, in their most basic form, assume default the first time that some credit indicator falls below a specified threshold value. In Merton's model (Merton (1974)) default occurs at the maturity date of debt provided the issuer's assets are less than the face value of maturing debt at that time.

Reduced-form models treat default as governed by a counting (jump) process coupled with an associated (possibly state-dependent) intensity process and, thus, whether or not an issuer actually defaults is an unpredictable event.

Several works deal with the empirical estimation of the structural models. Among others, Eom, Helwege, and Huang (2003) empirically test five structural models (Merton (1974), Geske (1977), Leland and Toft (1996), Longstaff and Schwartz (1995) and Collin-Dufresne, Goldstein, and Martin (2001)) of corporate bond pricing using data on the US market. They clearly show that all the five models considered have relevant spread prediction errors. In particular all the models tend to underestimate the spread of higher rated corporate bonds while they overestimate the spread of bonds which are considered riskier. A less structural approach has addressed the question of which variables are most correlated with the credit spread movements following a data-driven approach. In this framework Duffee (1999) investigates the effect of the term structure on callable and non callable credit spread.

This paper studies the determinants of credit spread changes in the Euro Corporate Bond Market. In particular, we are interested in understanding to what extent the implications for credit spread changes of structural credit risk models are verified in the context of the Euro corporate bond market.

Collin-Dufresne, Goldstein, and Martin (2001) show that variables postulated in the structural approach have a rather limited explanatory power. They consider other variables than those prescribed by the structural models in order to capture other effects such as the liquidity premium and the dynamic of interest rates. To this end they adopt a heterogenous parameter model for each issue and find out that the residuals from these regressions are highly cross correlated. A principal component analysis of the residuals shows that the first component is able to explain over 75 percent of the total variation of credit spreads. They also find that this large systematic component is not explained by several financial as well as macroeconomic variables. The authors conclude that the common systematic factor that drives the credit spread changes is a local domand/supply factor shock that is independent of traditional credit risk factors.

Elton, Gruber, Agrawal, and Mann (2001) move in a different direction. They point out that credit spread changes are determined not only by credit risk but also by risk premium. Credit spread changes can be easily viewed as an excess return of corporate bonds over treasury, i.e. risk free bond proxy. Therefore, they approach the problem in the framework of a traditional equity factor model to assess the influences of stock return common factors on credit spread.

So far, most of the empirical works on credit spreads deal with US data and relatively little is known about to the extent to which these results apply to the Euro market. Even though the empirical analysis of the US corporate bond market is an obvious reference, the European market is characterized by marked differences.

While in Europe the bond market is dominated by government bonds and bonds issued by the financial intermediaries, the bond market in the United States is dominated by the non financial corporate sector. In addition, municipal and agency bonds are major components of this market. Moreover, at least in Europe, it is difficult and costly to short sell corporate bonds (see Biais, Declerck, Dow, Portes, and von Thadden (2006)). This further reduces the liquidity of the corporate bond market. Annaert and Ceuster (2000), using data on aggregate index by rating categories and maturity buckets, stress that the European bond market shows broad similarities with the US market. In particular, the average spreads increase both by credit risk and by maturity and there is evidence of a strong correlation between credit spread and the determinants of interest rate term structure. However, their results are based on a rather limited time period when the Euro corporate bond market was not yet well developed.[1] Houweling, Mentink, and Vorst (2005) analyze the excess yield on corporate bonds. They use several proxies to test whether liquidity is priced in the euro-denominated corporate bond market. Under both the assumptions on constant and time-varying liquidity premium, they find strong evidence of priced liquidity. de Jong and Driessen (2005) consider liquidity proxies of equity markets and show that returns on corporate bonds are correlated with market-wide fluctuations in the liquidity of the equity market. They provide evidence that the European corporate bond excess holding returns have a significant exposure to liquidity risk and that a liquidity premium helps to explain part of the "credit spread puzzle", i.e. corporate bond yield spreads wider than what predicted by historical default losses.

We first analyze to what extent the European delta credit spreads are influenced by an underlying unobserved common factor which can be identified as a "market factor", as shown by Collin-Dufresne, Goldstein, and Martin (2001) for the American corporate bond market .

Adopting their specification we find that when we control also for the "market", the fitted residuals from the regressions of the delta credit spreads on the observed individual and common factors are still highly cross correlated. Thus suggesting the presence of a common systematic factor, which, obviously, cannot be identified with the market.

Moreover, the individual regressions show a substantial parameter heterogeneity across bonds. In general, the presence of unobserved factors, evident from the analysis of fitted residuals of both univariate regressions and the fixed effects panel data model, suggests

[1] They consider daily data ranging from March 1998 to May 1999.

the adoption of a consistent econometric estimation procedure. In fact, whenever an unobserved common factor structure exists the estimates of individual slope coefficients are inconsistent (see Bai (2006), Coakley, Fuertes, and Smith (2002), Coakley, Fuertes, and Smith (2006), Pesaran (2006)).
To this end the delta credit spreads are expressed as a function of both individual components and observed and unobserved common factors, where the former are linearly dependent on the latter. This amounts to have a heterogeneous model, i.e. varying slope coefficients, with a multifactor error structure. In this framework, studied in different papers (for instance Bai (2006), Coakley, Fuertes, and Smith (2002), Coakley, Fuertes, and Smith (2006), Pesaran (2006)), we can consistently estimate the effects of observed common and individual factors only if we take into account the relations between observed and unobserved factors, as shown in Pesaran (2006).
The estimation is based on a two-step procedure. First, the effects of observed factors are estimated using the Common Correlated Estimator, put forward by Pesaran (2006), providing a consistent estimate of the errors, which are supposed to be linearly dependent on a set of unobserved common factors, then the factors are estimated using a principal components analysis (see Bai (2003)), and eventually the model is reestimated under the assumption that the coefficients are random (Swamy (1970)). We show that this procedure yields consistent estimates, in average norm, of the unobserved factors. We also provide a Monte Carlo evidence.
We find that in general the variables suggested by the theory are both economically and statistically significant in explaining variations in individual issues' credit spreads. However, the factors predicted by the structural model are not as relevant as in the case of the US market.

There is strong evidence that one unobserved common factor has a relevant influence on the delta credit spreads. Even though the estimated factor is not identified we suspect that this is related to the market liquidity conditions. In fact, the European bonds seem to be mispriced as we show by comparing the actual to the fitted prices, using a smoothed cubic spline. Moreover the existing literature presents strong empirical evidence of a liquidity premium on the European market similar to what found for the US.

Finally, the paper is organized as follows. In section 2 we discuss the meaning of the credit spread changes. Section 3 introduces the structural credit risk models. In section 4 we present the individual and common factors used in the analysis. Preliminary empirical evidence is presented in section 5. The econometric model is introduced in section 6. Section 7 describes the data. Results are discussed in section 8. Section 9 concludes our findings.

2 Delta Credit Spread and Excess Returns

We define credit spread as the difference between the yield to maturity on a corporate bond and the yield to maturity on a government bond of the same maturity:

$$cs_t = c_t - g_t \tag{1}$$

where c_t is the redemption yield of a corporate bond at time t and g_t is the corresponding (i.e. with the same maturity) redemption yield on a government bond.[2]

The return on a coupon bond j for an holding period equal to one is given by

$$r_{j,t} = \frac{(P_{j,t} + C_{j,t}) - P_{j,t-1}}{P_{j,t-1}}$$

where $P_{j,t}$ is the gross price at time t for bond j and $(C_{j,t})$ is the interest or coupon payments of bond j at time t. We employ the approximation developed by Shiller (1979) from the first-order Taylor's approximation of the asset price as function of the corresponding yield:

$$r_{j,t} \cong -d_{j,t}(y_{j,t} - y_{j,t-1}) \tag{2}$$

where $y_{j,t}$ is the redemption yield[3] of bond j at time t and $d_{j,t}$ is the modified duration of bond j at time t.[4] Hence, the modified duration indicates the percentage change in the price of a bond for a given change in yield. Using expression (2), the difference between the return on the corporate bond and the government position can be written as

$$r_{c,t} - r_{g,t} \cong -d_{c,t}(c_t - c_{t-1}) + d_{g,t}(g_t - g_{t-1}) \tag{3}$$

where $r_{g,t}$ and $r_{c,t}$ are the return on the government and corporate bond, respectively.

We know that holding other factors constant, the lower the yield to maturity and the coupon rate the higher the duration. In general, the corporate bonds have a higher coupon rate and a higher yield than a government bond with the same maturity. Hence, the duration of government bonds can be thought of as the duration of corporate bonds plus, in general, a positive spread, $\gamma(t)$:

$$d_{g,t} = d_{c,t} + \gamma(t)$$

Then, expression (3) becomes:

$$r_{c,t} - r_{g,t} \cong -d_{c,t}\delta_t + \gamma(t)(g_t - g_{t-1}) \tag{4}$$

where δ_t, i.e. the *delta credit spread* is defined as:

$$\delta_t = cs_t - cs_{t-1} \tag{5}$$

[2]In particular, g_t is given by the redemption yield on the estimated euro government curve. See section 3.

[3]The redemption yield of bond j at time t, $y_{j,t}$, equals the internal rate of return that discounts its cash flows, including the interest or coupon payments $(C_{j,t})$ and the repayment of principal (PC_j), back to the bond's current price, i.e. $P_{j,t} = \sum_{\tau=t}^{T-1} C_{j,\tau}(1 + y_{j,t})^{-\tau} + (C_{j,T} + PC_j)(1 + y_{j,t})^{T}$.

[4]There are two measures of duration. Macauley duration (md) and modified duration (d). Macauley duration is the weighted average time to maturity of a bond, where the weights are given by the present values of the cash flows:

$$md_{j,t} = \sum_{\tau=t}^{T} \tau \times w_\tau = \sum_{\tau=t}^{T} \tau \times \frac{\frac{C_{j,\tau}}{(1+y_{j,t})^\tau}}{\sum_{\tau=t}^{T} \frac{C_{j,\tau}}{(1+y_{j,t})^\tau}}$$

The modified duration is simply the Macauley duration as defined above divided by $(1 + y_{j,t})$.

the second term on the RHS of (4) is negligible with respect to the excess return. Hence, the excess return of corporate bonds over government bonds is proportional to the change in credit spread:

$$r_{c,t} - r_{g,t} \cong -d_{c,t}\delta_t \tag{6}$$

The delta credit spread represents a proxy for corporate bond excess loss, that is, the return on a government bond minus the return on a corporate bond with the same maturity. Recently, large part of empirical analysis (in particular the practitioners) have changed their focus from bond yields to delta credit spreads. One example, as stressed by Collin-Dufresne, Goldstein, and Martin (2001), is represented by European mutual funds which invest both in corporate and government bonds. As a consequence, their portfolios are extremely sensitive to changes in credit spreads rather than changes in bond yields. Another example is represented by hedge funds trading strategy of taking highly leverage positions in corporate bonds while hedging away interest rate risk by shorting government bonds.

3 Structural models of credit risk

The issuer of a fixed-income security might default prior to the maturity date. This means that both the magnitude and the timing of payoffs to investors may be uncertain. How these default risks affect corporate bond pricing depends on how the default event is defined and how recovery in the event of a default is specified. The pricing models can be classified into *reduced form* models, those that are based on an assumed default intensity, and *structural* models, where there is an explicit characterization of the default event, like the first time that a firm's assets fall below the value of its liabilities (Duffie and Singleton, 2003).

In this paper we refer to the structural-model approach and to risk premia theory in order to identify the main factors (individual and common) that drive credit spread changes. The seminal papers of Black and Scholes (1973) and Merton (1974) introduced the first model of the structural-form approach. In the Black-Scholes-Merton model we may think of equity and debt as derivatives with respect to the total market value of the firm, and priced accordingly. We are in the setting of standard Black-Scholes model, i.e. a market with continuous trading which is frictionless and competitive.[5] The original owners of the firm choose a capital structure consisting of pure equity and of debt in the form of a single zero-coupon bond maturing at time T, of face value D. In the event that the total value V_T of the firm at maturity is less than the contractual payment D due on the debt, the firm defaults, giving its future cash flows, worth V_T, to debt holders. The debt can then be viewed as a difference between a riskless bond and a Black-Scholes price of a European put option on the firm's asset. The option representation of the bond price implies that:

- it is increasing in V.
- it is increasing in D.

[5] In detail, the agents are price-takers, there are no transactions costs, there is unlimited access to short selling and no indivisibilities of assets, and borrowing and lending through a money-market account can be done at the same riskless, compounded rate r.

- it is decreasing in the riskless interest rate.
- it is decreasing in time-to-maturity
- it is decreasing in the firm's value volatility.

The Black-Scholes-Merton model shows important drawbacks.[6] In particular, it mainly focuses on the value and the capital structure of the firm, which is a difficult process to represent. Besides that, the structural approach provides an intuitive framework to determine the main factors that drive credit spread changes. In the next section, we present the set of variables used in the analysis of the Euro corporate bond market, which are inspired by the structural-model approach (see Avramov, Jostova, and Philipov, 2007).

4 Individual and common factors

Our determinants of credit spread changes are inspired by structural default risk models. The contingent-claim approach views debt as a combination of a risk-free loan and a short put option on the firm. Variables governing the firm-value process affect default probabilities and recovery rates and ultimately drive credit spreads. Structural model variables typically include interest rates, term-structure slope, market return, market volatility, as well as firm leverage and volatility. In the following we present the variables we use in the empirical analysis that are supposed to affect credit spread changes. We distinguish between common and specific factors.

4.1 Common factors

1. *Changes in the government bond rate level.* This variable represents both a proxy for, the so called, *flight to quality* flows and a proxy for business cycle. From one side, a lower level of government rates implies a market preference for less risky asset, i.e. wider credit spreads. From the other side, lower rates also imply a higher loan demand which widens the credit spreads. Empirical evidence that a negative relationship exists between changes in credit spreads and interest rates has been shown by Longstaff and Schwartz (1995), Duffee (1998) and Collin-Dufresne, Goldstein, and Martin (2001). We use the DataStream's monthly series of the 10-year Benchmark German Treasury rates (denoted as *Gov*) to compute the monthly changes (denoted as *10Gov*), and the monthly variation of the squared 10-year Benchmark German Treasury rates ($10Gov^2$).

2. *Changes in the slope of the government yield curve.* This is a proxy of the movement in the supply and demand of government bonds. Hence a flat term structure of

[6] First, it requires inputs related to the value of firms that are often unavailable. Second, it allows default only at the maturity date of the bond. Third, it assumes independence between interest rates and credit risk. Last but not least, because it assumes that the value of the asset follows a geometric Brownian motion, the model implies that the default is predictable shortly before default. The first structural model has been widely improved by relaxing some of its restrictive assumptions, see, among others, Black and Cox (1976), Turnbull (1979), Leland (1994), Longstaff and Schwartz (1995), Briys and De-Varenne (1997) Collin-Dufresne and Goldstein (2001).

interest rates curve reduces the incentives to invest in the government sector and therefore causes a corporate spread widening. Duffee (1998) tests this relation for the US corporate bond market. Moreover, a steepening of the term-structure slope implies an increase in expected future spot rates, thereby reducing credit spreads.

The changes in the slope of the yield curve is given by the monthly changes in the difference between DataStream's 10-year and 2-year Benchmark German Treasury rates (denoted as *Slope*).

3. *Changes in the convexity of the government yield curve.* We also include the convexity of the government yield curve to capture potential non linear effects. This is calculated as the monthly changes in the difference between the 5-year German Treasury rate and the average of the 10-year and the 2-year Benchmark German Treasury rates (denoted as *Conv*).[7]

4. *Changes in liquidity.* Collin-Dufresne, Goldstein, and Martin (2001) stressed the fact that the corporate bond market tends to have relatively high transaction costs and low volume. These findings suggest checking for the existence of a liquidity premium. Given the fact that the corporate bond market it is an over-the-counter market is difficult to assess its liquidity unless we observe bid-ask spreads and/or volumes, which is not the case here. The standard measures of liquidity are unavailable in this case. Following Houweling, Mentink, and Vorst (2005), that scrutiny different liquidity proxies for the European corporate bond market, we consider several proxies to measure variations in liquidity: the monthly change in the five year Euro swap spread (denoted as *5dss*), the monthly variation in the number of issues of the corporate bonds included in the IBOXX Euro Bond Index, (*nofissue*). Furthermore, considering that the liquidity can be correlated with the return volatility we include the squared index monthly return (*ret2*) as suggested by Hong and Warga (2000).[8]

   Because of the strong link between swap and corporate market,[9] we expect that a change in the swap market liquidity would reflect a change in the same direction in the corporate market liquidity.[10] A decrease in the swap market liquidity (i.e. in the corporate market) implies a market preference for less risky assets. Hence we expect the factor loading of this liquidity proxy to be positive. Second, the issued amount of a bond is often assumed to give an indication of its liquidity. The higher the issued amount, the higher the liquidity of a bond. When the liquidity of the corporate bond market increases, corporate bond prices increase and credit spreads decrease. Hence, we expect a negative effect of this liquidity proxy on the delta credit spreads.[11] Since we do not observe the issued amount for each bond,

[7]The data source is DataStream.

[8]Though ARCH modeling can be adopted here to estimate the conditional variance as a measure of volatility there is no evidence of ARCH effects in the monthly index return.

[9]The issuers of corporate bonds typically fund on the swap market. Thus, if swap spreads widen, the long-term funding costs of corporate bonds issuers should increase, and investor demand for credit bonds should decrease. Assuming a constant supply of bonds, the decline in demand for credit products will cause prices to decline and the spread to Treasury to widen

[10]See Collin-Dufresne, Goldstein, and Martin (2001).

[11]Houweling, Mentink, and Vorst (2005) provide an extensive survey on both the theoretical structure and the empirical applications which consider the issued amount as liquidity proxy.

we consider a market liquidity measure such as the the monthly variation in the number of issues of the corporate bonds included in the IBOXX Euro Bond Index.

5. *Changes in the Business Climate.* Even if the probability of default remains constant for a firm, changes in credit spreads can occur due to changes in the expected recovery rate. The expected recovery rate in turn should be a function of the overall state business climate. We use stock indices return as proxies for the overall state of the economy and we expect that an increase in the index return reduces the credit spreads. We use the monthly Morgan Stanley Euro Index price return (denoted as *MSeuro*) as a proxy of the overall state of the economy.

6. *Credit Spread.* To investigate the presence of a mean-reverting behavior in credit spreads, we include the beginning-of-month level of credit spread (denoted as *Spread*). In case of a mean-reverting behavior this variable should contain information about the current month's change in credit spread.

7. *Changes in Credit Quality.* Changes in credit quality which also includes downgrading or upgrading in rating is part of credit risk. A general process of improvement or worsening in the credit quality should inversely move the credit spreads: a better credit quality reduces credit spread. We proxy the change in credit quality by monthly changes in rating downgrading (denoted as *Downg*) and upgrading (denoted as *Upg*) of the Merrill Lynch Global High Grade Corporate Index.[12]

4.2 Firm-specific factors

The firm-specific factors considered are:

1. *Mean and Standard Deviation of daily excess return of firm's equity.* These variables summarize the firm-level risk and return. Equity data reflect up-to-date information on firm value and should anticipate bond prices[13]. An increase in the equity daily excess return means a higher firm profitability. In line with the analysis of Kwan (1996) and Campbell and Taksler (2003) we expect stock returns to have a negative effects on credit spreads. We note also that previous studies of yield changes have often used the firm's equity return instead of changes in leverage as proxy for changes in the firm's health. On the other hand it is well known that the equity volatility of a firm increases its probability of default. Hence, a firm's volatility should drive up the yields on corporate bonds and widen the credit spreads. Moreover, Houweling, Mentink, and Vorst (2005) assume that bonds issued by companies whose equities are listed on a stock market are more liquid. Therefore, our sample should contain corporate bonds with higher liquidity and lower yields with respect to the full sample. Following Campbell and Taksler (2003) we match bond data with

[12] We take into account only Euro denominated bonds and the monthly changes are computed with respect to the index par amount. The data come from the Merrill Lynch Index Rating Migration Databook. This databook summarises relevant information on the composition of the main Merrill Lynch Corporate Bond Indices.

[13] Ederington, Yawitz, and Roberts (1987) claim that all data going into ratings prices should be anticipated by equity prices. Moreover they argue that investors fully anticipate rating changes which almost never affect bond returns.

equity data to explicitly evaluate the effects of equity volatility on corporate bond yield spreads. We consider only the corporate bonds issued by firms included in the Morgan Stanley World All Country Index[14]. To compute the daily excess return of each firm's equity we consider the Morgan Stanley Indices of the country where the stock is exchanged[15]. For each firm's equity we compute the mean (denoted as *Avgret*) and standard deviation (*Stdret*) of daily excess returns over the 180 days prior to (not including) the bond trade.

2. *Changes in Credit Market Factors.* We test whether credit spread changes depend on bond characteristics such as rating and the industrial sector. We also consider market factors for rating categories and industrial sectors. Each bond is assigned to one of the IBOXX sub-indices based on the bond's beginning of month rating or sector. We consider four rating categories, (AAA, AA, A, BBB) (denoted as *Dcsrat*) and three industrial sectors, (Industrial, Financial, Utility) and for each sub-index we consider the index monthly delta spread (*Dcsect*).

We do not use accounting variables to explain the credit spread changes. This choice is driven by two considerations. First, accounting data have in general either quarterly or yearly frequency. We think that interpolating the data does not provide so much information on credit spread changes. Second, most of the works which use accounting variables do not find any statistical evidence of their explanatory power and conclude that they can hardly explain the observed movements in credit spreads.
A detailed description of the data set is postponed until section 7.

5 Preliminary empirical evidence for the European corporate bond market

The important conclusions of Collin-Dufresne, Goldstein, and Martin (2001) for the US corporate bond market seem to be a logical starting point for any empirical study of the European corporate bond market. In this section, we try to replicate their analysis using data on Euro denominated corporate bonds. The goal is to understand to what extent the European market resembles the American market at least for what concerns the factors that affect the delta credit spreads.
They start from a simple model where the delta credit spread of each bond i at time t, y_{it}, depends on common observed factors, $\boldsymbol{d_t}$, and individual specific components, $\boldsymbol{x_{it}}$:

$$y_{it} \;=\; \boldsymbol{\alpha_i}'\boldsymbol{d_t} + \boldsymbol{\beta_i'}\boldsymbol{x_{it}} + e_{it} \qquad t = 1, \dots, T \qquad i = 1, \dots, T$$

In line with their analysis, we consider three different specifications. Nevertheless, we have to mention that for the European market we lack some of the data corresponding to the US market. However, we attempt to follow, as closely as possible, their specification

[14] The data source is DataStream.

[15] We end up considering the following Morgan Stanley Indices: Msci Emu, Msci Denmark, Msci Finland, Msci Norway, Msci Sweden, Msci Switzerland, Msci Uk, Msci Usa, Msci Canada, Msci Japan And Msci Hong Kong.

approach. They found that nearly half of the variation in spreads was unaccounted for by their regressors.

The first specification includes the following common factors ($\boldsymbol{d_t}$):

- the monthly change in the German government slope (*Slope*)
- the monthly change in the ten year German government bond yield-to-maturity (*10Gov*)
- the monthly change in the German convexity (*Conv*)
- the monthly return on Morgan Stanley Euro markets index (*MSeuro*)

and the individual factors ($\boldsymbol{x}_{it}$):

- the average of daily excess equity returns over the preceding 180 days (*Avgret*).
- the standard deviation of daily excess equity return over the preceding 180 days (*Stdret*).

The first model specification is:

$$\begin{aligned} y_{it} =& \alpha_{1i} + \alpha_{2i}\text{Slope}_t + \alpha_{3i}\text{10Gov}_t + \alpha_{4i}\text{Conv}_t + \alpha_{4i}\text{MSeuro}_t \\ &+ \beta_{1i}\text{Avgret}_{it} + \beta_{2i}\text{Stdret}_{it} + e_{it} \qquad t = 1, \dots, T \end{aligned} \tag{7}$$

The second specification includes additional explanatory variables to control for possibly omitted systematic common factors:

- the spread of the IBOXX euro corporate bond index at time $t-1$ (*Spread*)
- the five year delta swap spread (*5dss*)
- the monthly variation in the total issued amount of the IBOXX index (*Totaos*)
- the monthly variation in the square level of the ten year German government benchmark yield-to-maturity. ($10Gov^2$)
- the level of the ten year German government benchmark yield-to-maturity at time $t-1$ (*Gov*)
- monthly variation in upgraded Euro corporate bonds (*Upg*)
- monthly variation in downgraded Euro corporate bonds (*Downg*)
- the level of VDAX index at time $t-1$ (*Vdax*)
- monthly variation in the VDAX index (*Dvdax*)

the VDAX index is a volatility index of the DAX options traded at the Eurex. The five year delta swap spread ($5dss_t$) and the monthly variation in the total issued amount of the IBOXX index ($Totaos_t$) are two liquidity proxies. The second model specification is:

$$\begin{aligned} y_{it} =& \alpha_{1i} + \alpha_{2i}\text{Slope}_t + \alpha_{3i}10\text{Gov}_t + \alpha_{4i}\text{Conv}_t + \alpha_{4i}\text{MSeuro}_t \\ &+ \alpha_{5i}\text{Spread}_{t-1} + \alpha_{6i}5\text{dss}_t + \alpha_{5i}\text{Totaos}_t + \alpha_{6i}10\text{Gov}_t^2 + \alpha_{6i}\text{Gov}_{t-1} \\ &+ \alpha_{7i}\text{Upg}_t + \alpha_{8i}\text{Downg}_t + \alpha_{9i}\text{VDAX}_{t-1} + \alpha_{10i}\text{DVDAX}_t \\ &+ \beta_{1i}\text{Avgret}_{it} + \beta_{2i}\text{Stdret}_{it} + e_{it} \qquad t = 1, \ldots, T \qquad i = 1, \ldots, I \end{aligned} \tag{8}$$

Finally, the third specification adds to the observed common factors what Collin-Dufresne, Goldstein, and Martin (2001) call a "market factor" for the corporate bond market, that is the monthly change in the IBOXX BBB Index credit spread, $Iboxxbbb_t$.

Each specification includes the intercept and is estimated by OLS. We aggregate the bonds by maturity buckets, rating categories and industrial sectors. The individual regressions show a considerable parameter heterogeneity. [16]

Table 1 reports the average adjusted R^2 of the specifications above. We observe that there are some differences with the results obtained by Collin-Dufresne, Goldstein, and Martin (2001) for the US market. Most of the explanatory variables have some ability to explain the delta credit spreads and most of the estimated parameters are in line with the predictions of economic theory, however the explanatory power of all the specifications put together is slightly lower than that found for the US market. Collin-Dufresne, Goldstein, and Martin (2001) find an average adjusted R^2 of 21, 35 and 55 percent for their specifications while we find overall an adjusted R^2 of 18, 35and 45 percent, respectively (Table 1).

Collin-Dufresne, Goldstein, and Martin (2001) found that the unexplained component of the movement in credit spread changes can be ascribed to the presence of a single common factor. In fact, looking at the cross-section correlation in the residuals of the three specifications, they showed that controlling for the influence of a market factor dramatically reduces the correlation among the fitted residuals. In particular, the percentage of the total residual variance explained by the first principal component drops from about 76% to about 40%. They conclude that the "market factor" can be identified as a supply/demand shock.

We replicate their analysis in the case of the European market. We divide the residuals into nine bins, determined by three maturity groups (Short term: $<$ 4 years, Medium term: $\geq$ 4 years and $<$ 10 years, Long term: $\geq$ 10 years) and three industrial sectors (Financial, Industrial and Utility).[17] The estimates of individual bond regressions suggest that the estimated parameters are characterized by a heterogeneity, both at the individual bond level and at bin level. [18]

Table 2 shows that the first component explains a relevant part of the variability in the residuals (64.9%, 56.4%, and 53.8%, for the first, second and third specification respectively) for all the specifications considered while the analysis of the US market

---

[16]For ease of exposition we do not report the individual regressions results, but they are available upon request from the authors.

[17]The results are very similar when we divide the residuals into bins based on maturity and rating.

[18]On the contrary for the US market, Elton, Gruber, Agrawal, and Mann (2001) show that there is a substantial homogeneity in the estimated parameters.

shows that when a "market factor" is added the first component accounts for a small fraction of the remaining variation. In our case, the residuals' variability is still high (53.8%) even when we introduce a proxy for the market into the regression (specification three). Moreover Collin-Dufresne, Goldstein, and Martin (2001) show that in the first two specifications the first principal component can be seen as an equally weighted portfolio across the categories used to build the bins. On the contrary, the empirical evidence for the European market does not allow for the same conclusions. In the European case, there can be potentially more than one unobserved factor which influences the variation in credit spreads. Moreover, this analysis does not provide any clear clue to the identification of the unobserved factors.
While Collin-Dufresne, Goldstein, and Martin (2001) use the evidence on the effects of the inclusion of the "market factor" to support the presence of an unobserved common factor, we cope with this problem in a different way.

From the average correlation of the corporate bond delta spread, shown in Panel B of Table 3, it is evident that the delta credit spreads are cross correlated. The first two principal components of delta credit spreads account for 62% of the total variance and the correlation of the residuals of the second specification (8) is quite high. This seems to confirm the presence of a relevant cross-section dependence.

Table 3 also reports the test statistic for cross dependence by Pesaran (2004).[19] The hypothesis that the residual credit spread changes are cross sectionally independent is strongly rejected. All this suggests the adoption of a panel-data approach to the analysis of delta-credit spread cross-section dependence that evolves over time. Table 3 reports the average correlations of fitted residuals from a fixed effect model of third specification (equation (8) augmented with the Iboxx index).

Again, the average cross-section correlation of fixed effects fitted residuals of the third specification is far from negligible. This suggests that there could be some omitted explanatory variables. This is also the conclusion of Collin-Dufresne, Goldstein, and Martin (2001) for the American market. They remark that in this case these must be found among non-firm-specific factors.

6 Econometric Model

The results in the previous section suggest that a suitable setup for modeling the changes in delta credit spreads is a panel data model with an unobserved common factor structure.

[19]Pesaran (2004) proposes a test for cross-section dependence based on a simple average of the all pair-wise correlation coefficients of the Ordinary Least Square (OLS) residuals from the individual regressions in the panel. This test is applicable to a variety of panel data models and despite of the Breusch and Pagan Lagrange Multiplier test, it can be used when the cross section dimension is large relatively to the time series dimension. The Cross Section Dependence statistic (CD stat) is computed as:

$$CD = \sqrt{\frac{2T}{I(I-1)}} \left(\sum_{i=1}^{I-1} \sum_{j=i+1}^{I} \widehat{\rho}_{ij} \right)$$

where $\widehat{\rho}_{ij}$ is the sample estimate of the pair-wise correlation of the residuals ($\widehat{e}_i$ and $\widehat{e}_j$). Under the null hypothesis of no cross section dependence the CD statistic is distributed (as I and $T \rightarrow \infty$ with no particular order) as a standard normal distribution.

Kapetanios and Pesaran (2007) extend the standard asset return equations, routinely estimated in the finance literature, allowing for unobserved and observed factors. We adopt the same approach. In particular, following Pesaran (2006), we consider a linear heterogeneous panel data model where y_{it} is the observation on the delta credit spread at time t for the i^{th} issue for $i = 1, 2, \ldots, I$ and $t = 1, 2, \ldots, T$:

$$\begin{aligned} y_{it} &= \boldsymbol{\alpha_i}'\boldsymbol{d_t} + \boldsymbol{\beta_i}'\boldsymbol{x_{it}} + e_{it} \\ e_{it} &= \boldsymbol{\gamma_i}'\boldsymbol{f_t} + \epsilon_{it} \end{aligned} \tag{9}$$

where $\boldsymbol{d_t}$ is a $n \times 1$ vector of observed common effects, $\boldsymbol{x_{it}}$ is a $k \times 1$ vector of observed individual specific regressors, $\boldsymbol{f_t}$ is the $m \times 1$ vector of unobserved common factors and ϵ_{it} are the idiosyncratic errors assumed to be independently distributed of $(\boldsymbol{d_t}, \boldsymbol{x_{it}})$.

To allow for correlation between $\boldsymbol{f_t}$ and $\boldsymbol{x_{it}}$ we suppose that the individual specific factors are correlated with common (observed and unobserved) factors through:

$$\boldsymbol{x_{it}} = \boldsymbol{A_i}'\boldsymbol{d}_t + \boldsymbol{\Lambda_i}'\boldsymbol{f_t} + \boldsymbol{v_{it}} \tag{10}$$

where $\boldsymbol{v_{it}}$ are the specific components of $\boldsymbol{x_{it}}$ distributed independently of the common effects and across i. Following Pesaran (2006) we can combine the expressions (9) and (10) in a system

$$\boldsymbol{z_{it}} = \begin{bmatrix} y_{it} \\ \boldsymbol{x_{it}} \end{bmatrix} = \boldsymbol{B}_i'\boldsymbol{d}_t + \boldsymbol{C}_i'\boldsymbol{f}_t + \boldsymbol{u}_{it}$$

where

$$\mathbf{u}_{it} = \begin{bmatrix} \epsilon_{it} + \boldsymbol{\beta}_i'\boldsymbol{v_{it}} \\ \boldsymbol{x_{it}} \end{bmatrix}$$

$$\boldsymbol{B}_i = \begin{bmatrix} \boldsymbol{\alpha_i} & \boldsymbol{A}_i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \boldsymbol{\beta_i} & \mathbf{I}_k \end{bmatrix} \qquad \boldsymbol{C}_i = \begin{bmatrix} \boldsymbol{\gamma_i} & \boldsymbol{\Lambda_i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \boldsymbol{\beta_i} & \mathbf{I}_k \end{bmatrix}$$

We can rewrite (9) as a T-dimension system:

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{D\alpha}_i + \boldsymbol{X}_i\boldsymbol{\beta}_i + \boldsymbol{e}_i \\ &= \boldsymbol{D\alpha}_i + \boldsymbol{X}_i\boldsymbol{\beta}_i + \boldsymbol{F\gamma}_i + \boldsymbol{\epsilon}_i \end{aligned}$$

where $\boldsymbol{D}$ is a $(T \times n)$ matrix, $\boldsymbol{X}$ is $(T \times k)$, $\boldsymbol{F}$ is $(T \times m)$, i.e. $\boldsymbol{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_T)'$, and $\boldsymbol{\epsilon_i} = (\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{iT})'$ is $(T \times 1)$.

Assumption 1 *The idiosyncratic errors $\boldsymbol{\epsilon}_t$ are supposed to follow a zero-mean process:*

1. $E(\epsilon_{it}) = 0$, $E(\epsilon_{it}^2) = \sigma_i^2$.
2. $E(\epsilon_{it}\epsilon_{jt}) = \sigma_{ij}$ *with* $\sum_{i=1}^{I} |\sigma_{ij}| \leq M$ *for all* (i, j)*, where M is a positive constant.*
3. $p\lim_{I\to\infty} \frac{1}{I}\sum_{i=1}^{I} \boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i' = \boldsymbol{\Omega}$ $(T \times T)$ *with the largest eigenvalue bounded uniformly in i and T.*

Moreover the idiosyncratic errors ϵ_{it} are assumed to be independently distributed of $(\boldsymbol{d}_t, \boldsymbol{x}_{it})$.

Assumption 2 *The rank of $(\boldsymbol{\gamma}_i \quad \boldsymbol{\Lambda}_i)$ is m.*

Assumption 3 *The observed and unobserved common factors are assumed to be orthogonal,* $E(\boldsymbol{f}_t \boldsymbol{d}_t') = \mathbf{0}$, $\forall t$, *They are covariance-stationary with absolute summable autocovariances, and distributed independently of the individual error* ϵ_{is} *and* $\boldsymbol{v_{is}}$, *for all* i, t *and* s.

Assumption 4 $p\lim \frac{1}{T}\sum_t \boldsymbol{f_t f_t}' = \boldsymbol{\Sigma}_F$, *where* $\boldsymbol{\Sigma}_F$ *is a* $(m \times m)$ *positive definite matrix.*

Assumption 5 $p\lim \frac{1}{T}\sum_t \boldsymbol{d_t d_t}' = \boldsymbol{\Sigma}_D$, *where* $\boldsymbol{\Sigma}_D$ *is a* $(n \times n)$ *positive definite matrix.*

Assumption 6 $p\lim p\lim \frac{1}{T}\sum_t \boldsymbol{d_t x_{it}}' = \boldsymbol{\Sigma}_{DX_i}$ *where* $\boldsymbol{\Sigma}_{DX_i}$ *is a finite matrix.*

Assumption 7 *Factor loadings:* $\|\boldsymbol{\gamma}_i\| \leq \overline{\gamma} < \infty$ *and* $\|\boldsymbol{\Gamma}'\boldsymbol{\Gamma}/I - \boldsymbol{\Sigma}_\Gamma\| \rightarrow \mathbf{0}$ *where* $\boldsymbol{\Sigma}_\Gamma$ *is a* $(m \times m)$ *positive definite matrix (Bai (2003) and Bai and Ng (2002)). The factor loadings are treated as parameters.*

Assumption 1 concerns the idiosyncratic errors. We assume that the errors are cross-sectionally correlated but constant in time, i.e. stationary. In Assumption 2 and 7, like in Bai (2003) and Bai and Ng (2002), the factors loadings are considered non random but as shown in Bai (2006) we can extend our analysis to include random factor loadings, provided they are independent of the factors and idiosyncratic errors. Thus the factors are treated as parameters.
Pesaran (2006) put forward, using cross section averages of y_{it} and $\boldsymbol{x_{it}}$ as proxies for the latent factors, $\boldsymbol{f_t}$, a consistent estimator for $\boldsymbol{\beta_i}$. The basic idea behind the proposed estimation procedure, the Common Correlated Effects (CCE) estimator, is to filter the individual specific regressors by means of cross section aggregates in such a way that (as $I \rightarrow \infty$) the differential effects of unobserved common factors are eliminated asymptotically.

6.1 Estimation

We are interested in understanding the impact of unobserved factors on the observed changes in credit spreads. To this end we propose an estimation procedure which is articulated in two steps. First, the individual slope coefficients are estimated by CCE estimator of Pesaran (2006), then a consistent (in average norm) principal components estimation of the unobserved factors is obtained. The number of factors is assumed to be unknown but fixed. Second, we estimate $(\boldsymbol{\alpha_i},\ \boldsymbol{\beta_i},\ \boldsymbol{\gamma_i})$ in a random coefficient model.
The CCE estimator is given by augmenting the OLS regression of y_{it} on $\boldsymbol{x_{it}}$ and $\boldsymbol{d_t}$ with the cross-section averages $\overline{\boldsymbol{z}}_{\boldsymbol{t}} = \frac{1}{I}\Sigma_{i=1}^{I}\boldsymbol{z_{it}}$:

$$\widehat{\boldsymbol{\beta}}_i = (\boldsymbol{X}_i'\overline{\boldsymbol{M}}\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i'\overline{\boldsymbol{M}}\boldsymbol{y}_i \tag{11}$$

with

$$\overline{\boldsymbol{M}} = \boldsymbol{I}_T - \overline{\boldsymbol{H}}(\overline{\boldsymbol{H}}'\overline{\boldsymbol{H}})^{-1}\overline{\boldsymbol{H}}'$$

$\overline{\boldsymbol{H}} = (\boldsymbol{D}, \overline{\boldsymbol{Z}})$, $\boldsymbol{D}$ and $\overline{\boldsymbol{Z}}$ being, respectively, the $T \times n$ and $T \times (k+1)$ matrices of observations on $\boldsymbol{d}_t$ and $\overline{\boldsymbol{z}}_t$. Although $\overline{\boldsymbol{y}}_t$ and ϵ_{it} are not independent (i.e. endogeneity bias), their correlation goes to zero as $I \rightarrow \infty$.

We estimate the model (9)-(10) with the hypothesis that the observed common factors are uncorrelated with unobserved ones, i.e. $E(\boldsymbol{f_t d_t}') = \mathbf{0}$, $\forall t$. In order to deal with error cross-section dependence due to unobserved common factors we adopt the following procedure:

1. we consistently estimate the slope parameter $\widehat{\boldsymbol{\beta}}_i$ by means of the CCE estimator of equation (11), based on an estimate of $\boldsymbol{f_t}$ by means of cross-section averages, $\overline{\boldsymbol{z}}_t$, and $\boldsymbol{d}_t$.

2. for $i = 1, \ldots, I$ we estimate the residuals as:

$$\widehat{\boldsymbol{e}}_i = \overline{\boldsymbol{M}}_d(\boldsymbol{y_i} - \boldsymbol{X_i}\widehat{\boldsymbol{\beta}}_i) \tag{12}$$

where $\overline{\boldsymbol{M}}_d$ is given by

$$\overline{\boldsymbol{M}}_d = \boldsymbol{I_T} - \boldsymbol{D(D'D)^{-1}D'}$$

the presence of unobserved common factors correlated with the individual specific regressors do not cause the inconsistency of the parameter estimates of the observed common effects part ($\boldsymbol{\alpha}_i$)

Proposition 6.1 *Under Assumptions 1,3,4,5,6 $\boldsymbol{\alpha}_i$ is consistently estimated by*

$$\widehat{\boldsymbol{\alpha}}_i = \boldsymbol{(D'D)^{-1}D'}(\boldsymbol{y_i} - \boldsymbol{X_i}\widehat{\boldsymbol{\beta}}_i).$$

Proof. See Appendix B.
Proposition 6.1 implies that the fitted residuals $\widehat{\boldsymbol{e}}_i$ are consistent for $\boldsymbol{e}_i$.

3. The unobserved common factors are estimated, up to a non-singular transformation (i.e. *rotation indeterminacy*), by the method of least squares. This corresponds to minimize the following objective function

$$tr[(\widehat{\boldsymbol{E}} - \boldsymbol{F}\boldsymbol{\Gamma}')(\widehat{\boldsymbol{E}} - \boldsymbol{F}\boldsymbol{\Gamma}')']$$

where $\widehat{\boldsymbol{E}}$ is the $(T \times I)$ matrix: $\widehat{\boldsymbol{E}} = (\widehat{\boldsymbol{e}}_1, \widehat{\boldsymbol{e}}_2, \ldots, \widehat{\boldsymbol{e}}_I)$ and $\boldsymbol{\Gamma}' = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_I)$ is $(m \times I)$. Concentrating out $\boldsymbol{\Gamma} = \widehat{\boldsymbol{E}}'\boldsymbol{F}(\boldsymbol{F}'\boldsymbol{F})^{-1} = \widehat{\boldsymbol{E}}'\boldsymbol{F}/T$ and using the normalization $\boldsymbol{F}'\boldsymbol{F}/T = \boldsymbol{I}_m$ the objective function becomes

$$tr(\widehat{\boldsymbol{E}}'\boldsymbol{M}_F\widehat{\boldsymbol{E}}) = tr(\widehat{\boldsymbol{E}}'\widehat{\boldsymbol{E}}) - tr(\boldsymbol{F}'\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'\boldsymbol{F})/T$$

Therefore minimizing with respect to $\boldsymbol{F}$ is equivalent to maximize $tr[\boldsymbol{F}'(\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}')\boldsymbol{F}]$. The estimator of $\boldsymbol{F}$ is equal to the first J eigenvectors associated with the first J largest eigenvalues of the matrix $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'$. In order to consistently estimate the number of factors we make use of the information criteria proposed by Bai and Ng (2002). Thus by the definition of eigenvalues and eigenvectors $\widehat{\boldsymbol{F}}$ satisfies

$$\left[\frac{1}{IT}\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'\right]\widehat{\boldsymbol{F}} = \widehat{\boldsymbol{F}}\boldsymbol{V}_{IT}$$

where $\boldsymbol{V}_{IT}$ is a diagonal matrix which consists of the J largest eigenvalues of $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'$ arranged in decreasing order.[20] When we look at the asymptotic characteristics of the estimated factors we assume that m is known.[21] The following proposition shows the average norm consistency of $\widehat{\boldsymbol{F}}$ for $\boldsymbol{F}$.

Proposition 6.2 *Suppose assumptions 1-7 hold. Let* $\boldsymbol{G} = (\boldsymbol{\Gamma}'\boldsymbol{\Gamma}/I)(\boldsymbol{F}'\widehat{\boldsymbol{F}}/T)\boldsymbol{V}_{IT}^{-1}$. *Then* $\boldsymbol{G}$ *is an* $(m \times m)$ *invertible matrix, and*

$$\frac{1}{T}\|\widehat{\boldsymbol{F}} - \boldsymbol{F}\boldsymbol{G}\|^2 = \frac{1}{T}\sum_{t=1}^{T}\|\widehat{\boldsymbol{f}}_t - \boldsymbol{G}'\boldsymbol{f}_t\|^2 = O_p(1/\min(I,T))$$

Proof. See Appendix B.

In Appendix C we show, by means of a Monte Carlo simulation, the small sample properties of the estimated factors, using the least squares principal components estimator.

4. Finally, $\widehat{\boldsymbol{f}}_{\boldsymbol{t}}$ are used as regressors in the model, Bai (2003) shows that as long as $\sqrt{T}/I \to 0$ the error in the estimated factor is negligible, and for large I, $\boldsymbol{f}_{\boldsymbol{t}}$ can be treated as known. The model becomes:

$$y_{it} = \boldsymbol{\alpha_i}'\boldsymbol{d_t} + \boldsymbol{\beta_i}'\boldsymbol{x_{it}} + \boldsymbol{\gamma_i}'\widehat{\boldsymbol{f}}_{\boldsymbol{t}} + \varsigma_{it} \tag{13}$$

where ς_{it} is the idiosyncratic error and $\widehat{\boldsymbol{f}}_{\boldsymbol{t}}$[22] is the $(J \times 1)$ vector of the first J principal components of $\widehat{\boldsymbol{\Sigma}}$. We augment the model by the estimated factors since we are interested in evaluating their impact. We assume a Swamy random coefficient model, that is:

$$\begin{aligned} \boldsymbol{\beta}_i &= \boldsymbol{\beta} + \boldsymbol{\eta}_i \quad & \boldsymbol{\eta}_i &\sim \text{i.i.d.}(\boldsymbol{0}, \boldsymbol{\Omega}_\eta) \\ \boldsymbol{\alpha}_i &= \boldsymbol{\alpha} + \boldsymbol{\xi}_i \quad & \boldsymbol{\xi}_i &\sim \text{i.i.d.}(\boldsymbol{0}, \boldsymbol{\Omega}_\xi) \\ \boldsymbol{\gamma}_i &= \boldsymbol{\gamma} + \boldsymbol{\nu}_i \quad & \boldsymbol{\nu}_i &\sim \text{i.i.d.}(\boldsymbol{0}, \boldsymbol{\Omega}_\nu) \end{aligned}$$

where $\boldsymbol{\eta}_i$, $\boldsymbol{\xi}_i$, and $\boldsymbol{\nu}_i$ are independent for all i. The parameters can be estimated either by feasible GLS or by Mean Group estimator (Pesaran and Smith, 1995); for T sufficiently large the two estimators are algebraically equivalent (Hsiao and Pesaran, 2008). Due to the lack of space we only report the estimates obtained with Feasible GLS estimator. As expected, the estimates based on the Mean Group are very close to GLS estimates.[23]

[20] The scaling by IT does not affect $\widehat{\boldsymbol{F}}$.

[21] Their asymptotic distributions are not affected when the number of factors is unknown and is estimated (see Bai (2003)).

[22] Or any linear combination of them, i.e. $\boldsymbol{H}\widehat{\boldsymbol{f}}_{\boldsymbol{t}}$, where $\boldsymbol{H}$ is an invertible matrix such that $\widehat{\boldsymbol{f}}_{\boldsymbol{t}}$ is an estimator of $\boldsymbol{H}\boldsymbol{f}_{\boldsymbol{t}}$ and $\boldsymbol{H}^{-1}\widehat{\gamma}_i$ is an estimator of $\boldsymbol{\gamma_i}$.

[23] They are available upon request from the authors.

7 Data Description

7.1 Data

Our corporate bond data are extracted from the IBOXX Euro Bond Index. This index is issued by seven major investment banks[24]. Each bank is due to buy and sell every single asset belonging to the index. The index bond prices are determined by the following criteria. First, the highest and the lowest prices are excluded and then the price is given by the average of the other five prices. Moreover each asset included should have at least 500 million Euros of amount outstanding and its time to maturity should be greater than one year. Such criteria should be guaranteed to deal with tradable and liquid assets. In this way we try to reduce the liquidity premium of the Euro corporate market.

The IBOXX database[25] contains issue- and issuer-specific variables such as callability, maturity, coupon, industrial sector, rating, subordination level, issuer country, duration and several measures of credit spread. The IBOXX Euro Index is composed both of Euro government bond and investment grade Euro corporate bonds. We consider only the Euro corporate bonds. We start considering monthly observations for the period from January 2000 to November 2004.

Because our goal is to explain the behavior of investment grade Euro corporate bonds we eliminate all the bonds downgraded to high yield debt.[26] The bonds under consideration have standard cashflows - fixed rate coupon and principal at maturity. We exclude all bonds not rated, step-up notes, floating rate debt and convertible bonds. We also exclude bonds with call options, put options or sinking fund provisions. Moreover, we require issuer with publicly traded stock in order to estimate equity volatility and equity excess return.

To match corporate bonds by corresponding stock we first match corporate ISIN by Bloomberg ISIN of the underline stock and we use the latter to extract equity data from DataStream. We also require that an issue have six months of stock price data prior to the bond trade.

Last, in order to undertake principal component analysis of the residuals we restricted our sample to a balanced panel. We only take into account only issues which have belonged uninterruptedly to the index from the last observation backward. We ended up with 207 bonds for 33 monthly observations.

We use the fitted government curve spread provided by the IBOXX database. This spread is equal to the difference between the yield to maturity of the corporate bond and the corresponding (i.e. with the same maturity) yield to maturity on the estimated Euro government curve[27].

Elton, Gruber, Agrawal, and Mann (2001) suggest the use of spot rates rather than yield to maturity because arbitrage arguments hold with spot rates. The procedure of Elton, Gruber, Agrawal, and Mann (2001) consists of computing the corporate spread

[24] ABN AMRO, Barclays Capital, BNP Paribas, Deutsche Bank, Dresdner Kleinwort Wasserstein, Morgan Stanley and UBS Investment Bank.

[25] The database has been built by the optimization group at Fideuram Investimenti SGR, Milan.

[26] During the sample period considered, 12 issuers and 24 issues downgraded to high yield bonds. 2 issuers and 4 issues defaulted.

[27] The Euro government curve is estimated by a cubic spline. Moreover, only German and French government bonds enter the term structure estimation process.

as the difference between the spot rate on corporate bonds in a particular rating class and the spot rate for Treasury bonds of the same maturity. Both zero curves are usually estimated by standard methods such as the Nelson-Siegel procedure or Spline functions.

Campbell and Taksler (2003) follow the procedure of Elton, Gruber, Agrawal, and Mann (2001) to eliminate coupon effects from corporate bond yields. First, they estimate the corporate bond spot curves for sector and credit rating. Then they use the zero-coupon curve to estimate the corporate bond prices. Last, for each bond, they obtain the redemption yield from the estimated prices. As a consequence of their analysis, Campbell and Taksler (2003) raise some doubts over the need to measure corporate bond yield spreads in relation to a zero-coupon curve. In fact even though their analysis make use of both "redemption yield spread" and "estimated redemption yield spread" they obtain very similar results.

The use of "estimated redemption yield spread" makes sense only if the approximated corporate bond prices are truly closed to the observed one. In general in the Euro bond market this is not the case. In fact whatever the interpolated technique used (Nelson-Siegel, Cubic Spline with 5 knots) the results are quite poor. In Appendix A, we present some evidence concerning the magnitude of the estimation errors of redemption yield spread based on estimated corporate spot rates.

Table 6 presents summary statistics on the bonds and issuers in the sample. Because of the reduction of the sample to match the equity data and to deal with a balanced panel data set one may wonder if these bonds are representative of the overall Euro corporate market. A comparison of our sample to all the noncallable and nonputable bonds included in the IBOXX index for the period considered (on average about 374 issues) suggests that they are very close. In Table 7 bonds in the sample and bonds included in the IBOXX index are compared. The two samples have very similar distribution across credit ratings (panel A) and industrial sectors (panel B). In Table 8 the distribution across the industries stresses the fact that the banks make up almost 28% of the entire sample (panel A). The distribution across maturity bucket of our sample has a slight tendency toward medium and short term bonds (panel B). Though the average bond maturity in our sample is very close to the average bond maturity of the full sample (5.66 in Table 6).

Although the criteria of the IBOXX index should guarantee the liquidity of their components, Table 6 shows that the full sample contains outliers. The standard deviation of the full sample is twice our sample standard deviation. The maximum monthly credit spread change is about 466 basis points for our sample and 2530 basis points for the full sample. Therefore the extra return of a corporate bond with respect to a government bond can be 25 % in a month if we consider the full sample.

8 Results

Following the estimation procedure outlined in section 6, first, we consistently estimate the slope parameters $\boldsymbol{\beta}_i$ in the model (9)-(10) by means of the CCE estimator. Second, from the variance-covariance matrix of the consistently estimated residuals of equation (12), we obtain the principal components. Third, we select the common factors according to the information criteria IC_p of Bai and Ng (2002). The results (see Table 9) suggest the inclusion of just one factor in our model. Hence we augment our regressions with the first principal component. Finally the model is reestimated using the Swamy random coefficient estimator.

We estimate four different specifications, denoted in table 10 with A,B,C, and D. For each specification the unobserved factors are estimated from the residuals variance-covariance matrix. The variables and the predicted effects are presented in table (4). Specification A includes all the regressors (individual and observed common factors), B adds the estimated common factor. The specification C excludes the non significant all the liquidity proxies (the 5-year delta credit spread (*5dss*), the monthly variation in the number of issues of the corporate bonds included in the IBOXX Euro Bond Index (*nofissue*), and the squared index monthly return (*ret2*)); while D includes the estimated factor. Both Swamy's test of slope homogeneity as well as the version proposed by Pesaran and Yamagata (2007) where the cross section dimension could be large relative to the time series dimension, show that the null hypothesis of parameter constancy is strongly rejected.

In general, the variables suggested by the theory are both economically and statistically significant in explaining variations in individual issues' credit spreads.

1. *Changes in the government bond rate level.* The monthly variation of 10 year German government benchmark yield-to-maturity (*10gov*) is significant at the 5% significance level only when the estimated factor is included like in specification B.

2. *Changes in the slope of the government yield curve.* The slope in the German government yield curve is significant (10% significance level) and has a negative impact on the delta credit spread. That is when the curve is flattening this increases the credit spreads. This is in accordance with the findings of Duffee (1998) and Collin-Dufresne, Goldstein, and Martin (2001).

3. *Changes in the convexity of the government yield curve.* When liquidity proxies are included the convexity of the government yield curve has a negative and statistically significant effect on the change of delta credit spreads. Similarly, when we exclude the liquidity proxies but we include the estimated factor, which we interpret as liquidity factor, the convexity turns out to be significant. This captures possible nonlinearities in the relation between delta credit spreads and yield curve movements.

4. *Changes in liquidity.* The liquidity proxies included in specifications A and B are not significant at the 5% significance level. This suggests that these proxies are inadequate to catch the influence of liquidity conditions on the movements of delta credit spreads. However this does not exclude that liquidity is a relevant factor in explaining the movements in the excess returns over corporate bonds.

5. *Mean and Standard Deviation of daily excess return on equity.* The standard deviation of daily excess return (*stdret*) over the preceding 180 days[28], a volatility proxy, is not significant except for specification A, that is when the estimated common

[28]We also run the procedure to compute the standard deviation of equity daily excess return by using the preceding 90, 270 and 360 days prior the bond trade.

factor is not included in the regression. While the parameter of the variable *average of daily stock excess return* (*avgret*) is significantly negative in all but specification B.

6. *Changes in Credit Quality.* It is particularly interesting is that while the change in rating downgrade is always strongly significant the change in rating upgrade is not. However, when the estimated factor is present among the regressor, both are significant. It should be noticed that the effect of the monthly variation in the total amount outstanding of the upgraded bonds is smaller than the variation in the downgraded ones. This suggests that shocks in the credit market have an asymmetric effect on delta credit spreads.

7. *Changes in the Business Climate.* The Morgan Stanley Euro Index price return is significant in all the cases considered. The sign of the estimated coefficient is in accordance with the theory, which says that the market sentiment has a positive impact on the excess returns of corporate bonds. s

8. *Changes in Credit Market Factors.* The monthly delta spread of the IBOXX sub-indices, based on the bond's beginning of month rating classification, to which each bond issue in the sample is assigned, is strongly significant.[29] This could be interpreted as a market factor which explains a large part of the variation of the delta credit spreads. An increase of 100 basis points in this credit market factor augments the delta credit spreads by about 70 basis points.

9. *Credit Spread.* The initial credit spread level is negative and strongly significant in all the regressions, this is in accordance with a mean-reverting behavior of the delta credit spreads (the same is found by Collin-Dufresne, Goldstein, and Martin (2001)).

The regression results indicate that the inclusion of one estimated factor has an impact on the coefficient estimates, larger still for the common observed effects. This is probably related to the high correlation between the individual regressors and the common observed effects in our equations.

So far our analysis shows that in the Euro corporate bond market exists a systematic unobserved risk factor, unaccounted by the main common factors suggested by the theory.

The average correlation coefficient between actual and fitted values, computed for each bond, is about 0.53. From the empirical distribution 55% of estimated correlations are larger than 0.5.

Repeating the analysis carried out in section 5, we obtain that the first principal component of the residuals of model (B) explain only the 45.6% of the total variation (see Panel B in table 2). Thus suggesting that the inclusion of the estimated factor accounts for the common systematic unexplained component.

Our guess is that the estimated factor accounts for *latent* liquidity effects. This is reinforced by the observation that the liquidity proxies are unable to control for liquidity

[29] The results obtained with the IBOXX sub-indices for industrial sectors are less significant, so we choose to retain only the indexes for rating categories.

distortions. We consider different liquidity proxies and observe that no one has a significant effect. Moreover we argue that the liquidity distortions are possibly induced by the presence of imperfections in the Euro corporate bond market. This idea comes mainly from the evidence, stressed in section 7, that corporate bonds in the Euro market could be mispriced.[30]

In order to understand to what extent the factor is correlated with credit quality, we calculate the average partial correlation between the delta credit spreads and the estimated factor, after having controlled for all the explanatory variables contained in Table 10, specification A. Table 11 reports, as expected, that the average partial correlation increases as the credit rating decreases, i.e. as the liquidity conditions worsen. This is close to what has been found by de Jong and Driessen (2005). They show that both US and European corporate bonds are exposed to systematic liquidity shocks and that a liquidity risk premium helps to explain part of the credit spread. Importantly, the liquidity exposure is larger for lower-rated bonds.

Hence, we think that potentially an aggregate factor driving liquidity in the bond market could be correlated with the estimated common factor, in line with the findings of Collin-Dufresne, Goldstein, and Martin (2001).

9 Conclusion

In this paper we investigate the determinants of credit spread changes denominated in Euros. We point out that the change in credit spreads can be viewed as a proxy of the excess return of the corporate bonds over government bonds. For this reason we conduct our empirical analysis in a factor model framework. We also follow a data-driven approach recently developed for the US market which addresses the question of which variables are mostly correlated with the credit spread movements. Notable differences to the American market arise. First, the estimated parameters seem to be quite heterogeneous across bonds and bins used in the analysis. Second, the unexplained part of the movements in the delta credit spreads, which is due to unobserved common factors, is not correlated with the market. Nonetheless, we find highly cross-correlated residuals from the single-bond regressions. This suggests a heterogeneous panel data model with a multifactor error structure. In this setup we distinguish observed and unobserved common factors, and in order to consistently estimate the influences of individual factors influences we adopt a recently developed estimator (Pesaran (2006)). Starting from these estimates we show that the unobserved factors can be consistently (in average norm) estimated. Finally the effects of individual components and common observed and unobserved factors are reestimated assuming a random coefficient model. Overall our analysis shows that a systematic risk factor exists in the Euro corporate bond market that is independent of the main common factors predicted by the theory. The estimated factor can be thought of as capturing the liquidity bias, which in turn can be caused by the lack of a fully developed market. This interpretation also seems to be supported by the price misalignments found in the Euro corporate bond market.

[30] See Table 5.

References

Annaert, J., and M. D. Ceuster (2000): "Modelling European Credit Spreads," Discussion paper, Deloitte '&' Touche.

Avramov, D., G. Jostova, and A. Philipov (2007): "Understanding Changes in Corporate Credit Spreads," *Financial Analyst Journal*, 63(2), 90–105.

Bai, J. (2003): "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71(1), 135–171.

——— (2006): "Panel Data Models with Interactive Fixed Effects," .

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

Biais, B., F. Declerck, J. Dow, R. Portes, and E. von Thadden (2006): "European Corporate Bond Markets: Transparency, Liquidity, Efficiency," Discussion paper, Center for Economic Policy Research.

Black, F., and J. Cox (1976): "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, 31(2), 351–367.

Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy*, 81, 637–654.

Briys, E., and F. De-Varenne (1997): "Valuing Risky Fixed Rate Debt: An Extension," *Journal of Financial and Quantitative Analysis*, 32(2), 239–248.

Campbell, J., and G. Taksler (2003): "Equity volatility and Corporate Bond Yields," *The Journal of Finance*, 58(6), 2321–2349.

Coakley, J., A. Fuertes, and R. Smith (2002): "A Principal Component Approach to Cross-Section Dependence in Panels," Discussion paper, Birkbeck College.

——— (2006): "Unobserved Heterogeneity in Panel Time Series Models," *Computational Statistics and Data Analysis*, 50(9), 2361–2380.

Collin-Dufresne, P., and R. Goldstein (2001): "Do Credit Spreads Reflect Stationary Leverage Ratios?," *Journal of Finance*, 56(5), 1929–1957.

Collin-Dufresne, P., R. Goldstein, and J. Martin (2001): "The Determinants of Credit Spread Changes," *The Journal of Finance*, 56(6), 2177–2207.

de Jong, F., and J. Driessen (2005): "Liquidity Risk Premia in Corporate Bond Markets," mimeo.

Duffee, G. (1998): "The Relation between Treasury Yields and Corporate Bond Yield Spreads," *The Journal of Finance*, 53(6), 2225–2241.

——— (1999): "Estimating the Price of Default Risk," *Review of Financial Studies*, 12(1), 197–226.

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

Ederington, L., J. Yawitz, and B. Roberts (1987): "The Informational Content of Bond Rating," *Journal of Financial Research*, 10(3), 211–226.

Elton, E., M. Gruber, D. Agrawal, and C. Mann (2001): "Explaining the Rate Spread on Corporate Bonds," *The Journal of Finance*, 56(1), 241–71.

Eom, Y., J. Helwege, and J. Huang (2003): "Structural Models of Corporate Bond Pricing: an Empirical Analysis," *Review of Financial Studies*, 17(2), 499–544.

Geske, R. (1977): "The Valuation of Corporate Liabilities as Compound Options," *Journal of Financial and Quantitative Analysis*, 12(4), 541–552.

Hong, G., and A. Warga (2000): "An Empirical Study of Bond Market Transactions," *Financial Analysts Journal*, 56(2), 32–46.

Houweling, P., A. Mentink, and T. Vorst (2005): "Comparing Possible Proxies of Corporate Bond Liquidity," *Journal of Banking and Finance*, 29(6), 1331–1358.

Hsiao, C., and M. Pesaran (2008): "Random Coefficient Models," in *The Econometrics of Panel Data (Third Edition)*, ed. by L. Matyas, and P. Sevestre. Kluwer, Academic publishers.

Kapetanios, G., and M. Pesaran (2007): "Alternative Approaches to Estimation and Inference in Large Multifactor Panels: Small Sample Results with an Application to Modelling of Asset Returns," in *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, ed. by G. Phillips, and E. Tzavalis. Cambridge University Press, Cambridge.

Kwan, S. (1996): "Firm-Specific Information and the Correlation between Individual Stocks and Bonds," *Journal of Financial Economics*, 40(1), 63–80.

Lando, D. (2004): *Credit Risk Modeling*. Princenton University Press, Princenton.

Leland, H. (1994): "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, 49(4), 1213–1252.

Leland, H., and K. Toft (1996): "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads," *Journal of Finance*, 51, 987–1019.

Longstaff, F., and E. Schwartz (1995): "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," *The Journal of Finance*, 50, 789–819.

Merton, R. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance*, 29, 449–70.

Pesaran, M. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," mimeo.

——— (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

Pesaran, M., and R. Smith (1995): "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68(1), 79–113.

Pesaran, M., and T. Yamagata (2007): "Testing Slope Homogeneity In Large Panels," *Journal of Econometrics*, forthcoming.

Shiller, R. (1979): "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," *Journal of Political Economy*, 87, 1190–1219.

Swamy, P. (1970): "Efficient Inference in Random Coefficient Regression Model," *Econometrica*, 38, 311–323.

Turnbull, S. (1979): "Debt Capacity," *Journal of Finance*, 34(4), 931–940.

Van-Landschoot, A. (2003): "The term structure of credit spreads on euro corporate bonds," Discussion paper, Tilburg University, Center for Economic Research.

A Appendix

For each month, we estimate the zero coupon yield curves for each rating category[31] by a smoothed cubic spline. We then use these spot rates to discount the coupon corporate bonds cash flows and obtain the fitted price for each bond. We observe that the difference between actual prices and estimated prices are fairly consistent[32]. While for the government bonds the absolute average error is about 1 cent, for the corporate bonds it is about 20 cents for all rating categories.

Figure 1 and 2 show respectively the difference between the market prices and the fitted prices for the German government bonds and for the A rated euro corporate bonds at the same date. The average error over all the months considered is much greater than the average error found in other studies (see among others Elton, Gruber, Agrawal, and Mann (2001) and Campbell and Taksler (2003)) on the US corporate bond market but is comparable to the average error found for the Euro market according to the analysis conducted by Van-Landschoot (2003).

Van-Landschoot (2003) extend the Nelson-Siegel method to estimate the European term structure of credit spreads for different sub-rating categories. The analysis shows that the Nelson-Siegel method results in systematic errors that depend on liquidity, coupon and subcategories within the rating category (plus, flat, minus rating). Therefore Van-Landschoot (2003) extends the model with four additional factors in order to take into account the underline effects. The average yield error of the extended Nelson-Siegel model is fairly consistent. For example, the yield error for an A rated bond is close to 16 basis points for a two year maturity bond and 15 basis points for a five year maturity bond. Such yield errors cause the price errors to be fairly consistent. Table II illustrates this point. Table II shows the error between the observed market price and the estimated price for any two corporate bonds included in the IBOXX index on August 25^{th}, 2005. The first one is a two year maturity bond issued by Lehman Brothers while the second one is a five year maturity bond issued by France Telecom. Both bonds have an A rating. First, we compute the "estimated redemption yield" by adding the yield error to the observed redemption yield. Then we obtain the "estimated price" from the "estimated redemption yield". Table II shows that for the two bonds the error between the actual price and the estimated price is about 30 cents and 75 cents for 100 euros, respectively. The result does not significantly change for different redemption yields and coupon rates.

Actual and estimated prices can mainly differ because bonds within the same rating category are not homogeneous. Moreover, there are other possible reasons. First, credit ratings are revised infrequently and often with one lag. Second, corporate bonds could be mispriced. Finally the magnitude of fitted errors strongly suggests the use of the "observed redemption yield spread".

[31]We consider the following rating categories: AA, A and BBB.

[32]The euro government bonds considered here are those belong to the IBOXX Euro government bond Index.

B Appendix

Stacking the time series observations for i in (9):

$$\boldsymbol{y_i} = \boldsymbol{D}\boldsymbol{\alpha_i} + \boldsymbol{X}_i\boldsymbol{\beta_i} + \boldsymbol{F}\boldsymbol{\gamma_i} + \boldsymbol{\epsilon_i} \;\; i = 1\ldots, I \tag{14}$$

where

$$\begin{aligned}
\boldsymbol{y_i} &= (y_{i1}, y_{i2}, \ldots, y_{iT})' && (T \times 1)\\
\boldsymbol{X_i} &= (\boldsymbol{x_{i1}}, \boldsymbol{x_{i2}}, \ldots, \boldsymbol{x_{iT}})' && (T \times K)\\
\boldsymbol{D} &= (\boldsymbol{d_1}, \boldsymbol{d_2}, \ldots, \boldsymbol{d_T})' && (T \times N)\\
\boldsymbol{F} &= (\boldsymbol{f_1}, \boldsymbol{f_2}, \ldots, \boldsymbol{f_T})' && (T \times M)\\
\boldsymbol{\epsilon_i} &= (\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{iT})' && (T \times 1)
\end{aligned}$$

We first obtain the consistent CCE estimator of equation (11) and then we estimate $\boldsymbol{\alpha_i}$ in the following OLS regression:

$$\boldsymbol{y_i} - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \boldsymbol{D}\boldsymbol{\alpha_i} + \boldsymbol{\nu_i} \tag{15}$$

where ν_{it} are idiosyncratic errors.

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}} = (\boldsymbol{D}'\boldsymbol{D})^{-1}\boldsymbol{D}'(\boldsymbol{y_i} - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}) \tag{16}$$

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}} = \boldsymbol{\alpha_i} + (\boldsymbol{D}'\boldsymbol{D})^{-1}(\boldsymbol{D}'\boldsymbol{X}_i\boldsymbol{\beta_i} + \boldsymbol{D}'\boldsymbol{F}\boldsymbol{\gamma_i} + \boldsymbol{D}'\boldsymbol{\epsilon_i} - \boldsymbol{D}'\boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}) \tag{17}$$

Under assumptions 1,3,4,5,6 and given that $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$ is a consistent estimator of $\boldsymbol{\beta_i}$, then $\hat{\boldsymbol{\alpha}}_{\boldsymbol{i}} \xrightarrow{p} \boldsymbol{\alpha_i}$ as $T \to \infty$ for $i = 1, \ldots, I$.

Proof of Proposition 6.2. We use the following results throughout:

$$T^{-1}\|\boldsymbol{X}_i\|^2 = O_p(1)$$

or

$$T^{-1/2}\|\boldsymbol{X}_i\| = O_p(1)$$

and, averaging over i

$$\frac{1}{I}\sum_{i=1}^{I}\frac{\|\boldsymbol{X}_i\|^2}{T} = O_p(1).$$

Furthermore, from assumptions 4 and 5

$$T^{-1/2}\|\boldsymbol{F}\| = O_p(1)$$

$$T^{-1/2}\|\boldsymbol{D}\| = O_p(1)$$

and given the assumed normalization $\boldsymbol{F}'\boldsymbol{F}/T = \boldsymbol{I}_m$

$$T^{-1}\|\widehat{\boldsymbol{F}}\|^2 = m$$

$$T^{-1/2}\|\widehat{\boldsymbol{F}}\| = \sqrt{m}.$$

From

$$\left[\frac{1}{IT}\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}'\right]\widehat{\boldsymbol{F}} = \widehat{\boldsymbol{F}}\boldsymbol{V}_{IT}$$

with

$$\widehat{\boldsymbol{e}}_i = \overline{\boldsymbol{M}}_d(\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i) + \boldsymbol{F}\boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i)$$

$$\frac{1}{IT}\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}' = \frac{1}{IT}\sum_{i=1}^{I}(\overline{\boldsymbol{M}}_d(\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i) + \boldsymbol{F}\boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i))(\overline{\boldsymbol{M}}_d(\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i) + \boldsymbol{F}\boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i))'$$

we obtain

$$\begin{aligned}
\widehat{\boldsymbol{F}}\boldsymbol{V}_{IT} \quad = \quad & \frac{1}{IT}\sum_{i=1}^{I}\Big[\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d + \overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d \\
& + \overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d + \overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d \\
& + \overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d + \overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d + \overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d \\
& + \overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d + \overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\Big]\widehat{\boldsymbol{F}} \qquad (18)
\end{aligned}$$

where

$$\begin{aligned}
\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} &= \frac{1}{IT}\overline{\boldsymbol{M}}_d\boldsymbol{F}\left(\boldsymbol{\Gamma}'\boldsymbol{\Gamma}\right)\boldsymbol{F}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} \\
&= \boldsymbol{F}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right) - \overline{\boldsymbol{P}}_d\boldsymbol{F}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right) \\
&- \boldsymbol{F}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)\left(\frac{\boldsymbol{F}'\overline{\boldsymbol{P}}_d}{T}\right)\widehat{\boldsymbol{F}} + \overline{\boldsymbol{P}}_d\boldsymbol{F}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)\left(\frac{\boldsymbol{F}'\overline{\boldsymbol{P}}_d}{T}\right)\widehat{\boldsymbol{F}}.
\end{aligned}$$

Then the expression (18) can be rewritten as

$$\begin{aligned}
\widehat{\boldsymbol{F}}\boldsymbol{V}_{IT} - \boldsymbol{F}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right) &= \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} \\
&+ \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} + \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} \\
&+ \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} - \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{P}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}} \\
&- \frac{1}{IT}\sum_{i=1}^{I}\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{P}}_d\widehat{\boldsymbol{F}} + \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{P}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{P}}_d\widehat{\boldsymbol{F}} \\
&+ \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} + \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i - \widehat{\boldsymbol{\beta}}_i)\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} \\
&+ \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} + \frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}} \qquad (19)
\end{aligned}$$

From Proposition 1 in Bai (2003) the matrix $\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)$ is invertible. Then multiplying each side of (19) by $\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}$

$$\widehat{\boldsymbol{F}}\boldsymbol{V}_{IT}\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}-\boldsymbol{F} \;=\; [T1+T2+\ldots+T11]\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}$$

We have

$$\begin{aligned} T^{-1/2}\left\|\widehat{\boldsymbol{F}}\boldsymbol{V}_{IT}\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}-\boldsymbol{F}\right\| &\le T^{-1/2}\left(\|T1\|+\ldots+\|T11\|\right) \\ &\quad\cdot\left\|\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}\right\| \end{aligned} \tag{20}$$

Now we consider each term on the right. For the first term, given $\overline{\boldsymbol{M}}_d=\boldsymbol{I}-\overline{\boldsymbol{P}}_d$

$$\begin{aligned} &T^{-1/2}\left\|\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right\|\le \\ &\frac{1}{I}\sum_{i=1}^{I}\frac{\|\boldsymbol{X}_i\|^2}{T}\|\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i\|^2\frac{\|\widehat{\boldsymbol{F}}\|}{\sqrt{T}}+\frac{1}{I}\sum_{i=1}^{I}\frac{\|\boldsymbol{X}_i\|^2}{T}\|\overline{\boldsymbol{P}}_d\|^2\|\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i\|^2\frac{\|\widehat{\boldsymbol{F}}\|}{\sqrt{T}} \\ &-\frac{1}{I}\sum_{i=1}^{I}\frac{\|\boldsymbol{X}_i\|^2}{T}\|\overline{\boldsymbol{P}}_d\|\|\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i\|^2\frac{\|\widehat{\boldsymbol{F}}\|}{\sqrt{T}}-\frac{1}{I}\sum_{i=1}^{I}\frac{\|\boldsymbol{X}_i\|^2}{T}\|\overline{\boldsymbol{P}}_d\|\|\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i\|^2 \end{aligned}$$

From the rank assumption (Assumption 2) and Theorem 1 in Pesaran (2006) it follows that $\|\widehat{\boldsymbol{\beta}}_i-\boldsymbol{\beta}_i\|=O_p\left(\frac{1}{\sqrt{T}}\right)$. Moreover given that $\|\overline{\boldsymbol{P}}_d\|=\sqrt{n}$ and $T^{-1/2}\|\widehat{\boldsymbol{F}}\|^2=\sqrt{m}$, we obtain

$$T^{-1/2}\left\|\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{X_i}(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)'\boldsymbol{X_i}'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right\|\le O_p(\|\widehat{\boldsymbol{\beta}}_i\;-\boldsymbol{\beta}_i\|^2)=O_p\left(\frac{1}{T}\right).$$

Applying the same argument we can prove that the terms that depend on $\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i$ are each $O_p\left(\|\widehat{\boldsymbol{\beta}}_i-\boldsymbol{\beta}_i\|\right)=O_p\left(\frac{1}{\sqrt{T}}\right)$. Given Assumption 3 the terms $T5, T6, T7$ are $o_p(1)$. Let consider the following terms:

$$T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right)$$
$$T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right)$$
$$T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right)$$

The first is

$$T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right)=T^{-1/2}\frac{1}{TI}\sum_i\left(\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}-\overline{\boldsymbol{P}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right)+o_p(1)$$

$$T^{-1/2}\frac{1}{TI}\left\|\sum_i\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right\|\leq\frac{1}{\sqrt{I}}\left(\frac{1}{T}\sum_t\left\|\frac{1}{\sqrt{I}}\sum_i\boldsymbol{\epsilon}_{it}\boldsymbol{\gamma}_i'\right\|^2\right)^{1/2}\left\|\frac{1}{\sqrt{T}}\boldsymbol{F}\right\|\left\|\frac{1}{\sqrt{T}}\widehat{\boldsymbol{F}}\right\|$$

from Lemma 1.(ii) in Bai and Ng (2002)

$$\left(\frac{1}{T}\sum_t\left\|\frac{1}{\sqrt{I}}\sum_i\boldsymbol{\epsilon}_{it}\boldsymbol{\gamma}_i'\right\|^2\right)^{1/2}=O_p(1)$$

$$T^{-1/2}\frac{1}{TI}\left\|\sum_i\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right\|\leq\frac{1}{\sqrt{I}}O_p\left(1\right)O_p(1)\sqrt{m}=O_p\left(\frac{1}{\sqrt{I}}\right).$$

Analogously for $T^{-1/2}\frac{1}{TI}\sum_i\left(\overline{\boldsymbol{P}}_d\boldsymbol{\epsilon}_i\boldsymbol{\gamma}_i'\boldsymbol{F}'\widehat{\boldsymbol{F}}\right)$ and $T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\boldsymbol{F}\boldsymbol{\gamma}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right)$. The last term

$$T^{-1/2}\left(\frac{1}{IT}\sum_{i=1}^{I}\overline{\boldsymbol{M}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{M}}_d\widehat{\boldsymbol{F}}\right)=T^{-1/2}\frac{1}{IT}\sum_{i=1}^{I}\left(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\widehat{\boldsymbol{F}}+\overline{\boldsymbol{P}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{P}}_d\widehat{\boldsymbol{F}}-\overline{\boldsymbol{P}}_d\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\widehat{\boldsymbol{F}}-\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\overline{\boldsymbol{P}}_d\widehat{\boldsymbol{F}}\right)$$

now the first term

$$T^{-1/2}\frac{1}{IT}\left\|\sum_{i=1}^{I}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\widehat{\boldsymbol{F}}\right\|\leq\frac{1}{T}\left\|\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\right\|\left\|\frac{1}{\sqrt{T}}\widehat{\boldsymbol{F}}\right\|\leq\frac{1}{IT}\sum_{i=1}^{I}\|\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\|\cdot\left\|\frac{1}{\sqrt{T}}\widehat{\boldsymbol{F}}\right\|$$

but from Assumption 1.3

$$p\lim_{I\to\infty}\left\|\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\right\|=\left\|p\lim_{I\to\infty}\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\right\|=\|\boldsymbol{\Omega}\|=\lambda_{max}(\boldsymbol{\Omega})=O_p(1)$$

then

$$\frac{1}{T}\left\|\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'\right\|=O_p\left(\frac{1}{T}\right)$$

analogously for the remaining terms. Thus we obtain that,

$$T^{-1/2}\left\|\widehat{\boldsymbol{F}}\boldsymbol{V}_{IT}\left(\frac{\boldsymbol{F}'\widehat{\boldsymbol{F}}}{T}\right)^{-1}\left(\frac{\boldsymbol{\Gamma}'\boldsymbol{\Gamma}}{I}\right)^{-1}-\boldsymbol{F}\right\|=O_p\left(1/\min\left(\sqrt{T},\sqrt{I}\right)\right).$$

Assuming that $\boldsymbol{V}_{IT}$ is invertible the left side of 20 can be written as

$$T^{-1/2}\|\widehat{\boldsymbol{F}}-\boldsymbol{F}\boldsymbol{G}\|=O_p(1/\min(\sqrt{I},\sqrt{T}))$$

where $\boldsymbol{G}=(\boldsymbol{\Gamma}'\boldsymbol{\Gamma}/I)(\boldsymbol{F}'\widehat{\boldsymbol{F}}/T)\boldsymbol{V}_{IT}^{-1}$, taking the squares on each side gives the result.

C Appendix

In section 6 the estimation procedure of individual and observed and unobserved common factors effects is based on the estimate of the latter by principal components. We use simulations to assess the adequacy of the asymptotic result of proposition 6.2 in approximating the finite sample properties of the unobserved factors. We assume the following data generating process (DGP):

$$y_{it} = \alpha_{i1} + \alpha_{i2} d_{2t} + \beta_{i1} x_{1it} + \beta_{i2} x_{2it} + \gamma_i f_t + \epsilon_{it} \tag{21}$$

$$x_{1it} = a_{11} + a_{21} d_{2t} + \gamma_1 f_t + v_{1it} \tag{22}$$

$$x_{2it} = a_{12} + a_{22} d_{2t} + \gamma_2 f_t + v_{2it} \tag{23}$$

for $i = 1, \ldots, I$, and $t = 1, \ldots, T$. This DGP considers only two individual specific components, x_{1it} and x_{2it}, two observed common factors, d_{1t} and d_{2t}, and one unobserved common factor f_t, with $E(f_t d_{2t}) = 0, \ \forall t$. β_{i1} and β_{i2} are generated by

$$\beta_{ij} = 1 + \eta_{ij} \quad \text{for} \quad j = 1, 2 \quad \text{and} \quad \eta_{ij} \sim IIDN(0, 0.04)$$

and fixed across replications. $(\alpha_{i1}, \alpha_{i2})' \sim IIDN(\mathbf{0}, 0.04 \times \mathbf{I}_2)$. The parameters

$$A = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

and

$$\Gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$$

are generated as $vec(A) \sim IIDN(\mathbf{0}, 0.5 \times \mathbf{I}_4)$, and $IIDN(\mathbf{0}, 0.5 \times \mathbf{I}_2)$ respectively, and are not changed across replications.

The common factors and the individual specific errors are generated as independent stationary AR(1) processes with zero means and unit variances:

$$\begin{aligned} d_{1t} &= 1 \\ d_{2t} &= \rho_d d_{2,t-1} + \eta_{dt}, \quad t = -49, \ldots 1, \ldots, T. \\ d_{2,-50} &= 0 \end{aligned}$$

$$\begin{aligned} f_t &= \rho_f f_{t-1} + \eta_{ft}, \quad t = -49, \ldots 1, \ldots, T. \\ f_{-50} &= 0 \end{aligned}$$

for orthogonality between d_t and f_t, i.e. $E[f_t d_{2t}] = 0, \ \forall t$, we generate

$$\begin{bmatrix} \eta_{dt} \\ \eta_{ft} \end{bmatrix} \sim IIDN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1-\rho_d^2 & 0 \\ 0 & 1-\rho_f^2 \end{bmatrix} \right), \quad \rho_d = 0.5, \quad \rho_f = 0.5$$

$$\begin{aligned} v_{jit} &= \rho_{v_{ij}} v_{ji,t-1} + \nu_{ijt}, \quad t = -49, \ldots 1, \ldots, T. \\ \nu_{ijt} &\sim IIDN(0, 1-\rho_{v_{ij}}^2), \quad v_{ji,-50} = 0 \ \ j = 1, 2 \end{aligned}$$

and

$$\rho_{v_{ij}} \sim IIDU\,(0.05, 0.95)\,, \quad j = 1, 2.$$

The errors of y_{it} are generated as:

$$\begin{aligned} \epsilon_{it} &= \sigma_i \zeta_{it} \quad \text{for} \quad i = 1, \ldots, I \\ \sigma_i^2 &\sim IIDU\,(0.5, 1) \\ \zeta_{it} &\sim IIDN\,(0, 1) \end{aligned}$$

After having generated the data according to the DGP described above, first, we compute the CCE estimators of $\boldsymbol{\beta_i}$ in (21) and the resulting residuals, as computed in (12). Second, we estimate the factor as the first principal component of the estimated covariance matrix of residuals. Like in Bai (2003) to evaluate the estimate of a transformation of f_t, $\widehat{f_t}$, we compute the correlation coefficient between $\{f_t\}_{t=1}^{T}$ and $\{\widehat{f_t}\}_{t=1}^{T}$ for different cross-sectional and times series dimensions, for each Monte Carlo simulation. The table below reports the average correlation coefficient over 2000 repetitions for combinations of $T = 20, 30, 50, 100, 200$, and $I = 30, 100, 200$. The results suggest that the factor estimates are highly correlated with the unobserved factor. This seems to confirm the results in Bai (2003), obtained in a different context, that is as $\sqrt{T}/I \to 0$ the estimation error in the factor estimates is negligible.

Average correlation coefficients between f_t and $\widehat{f_t}$

I,T	*20*	*30*	*50*	*100*	*200*
30	0.9442	0.9608	0.9722	0.9806	0.9850
100	0.9565	0.9688	0.9797	0.9884	0.9925
200	0.9578	0.9716	0.9827	0.9901	0.9943

Small Sample Properties of Swamy estimator

		Bias				Rmse			
I	T	α_1	α_2	β_1	β_2	α_1	α_2	β_1	β_2
30	*20*	-0.007	0.004	0.001	0.000	0.402	0.310	0.059	0.059
100	*20*	0.009	-0.012	-0.001	-0.001	0.384	0.305	0.035	0.035
200	*20*	-0.008	0.007	0.000	0.000	0.386	0.301	0.023	0.023
30	*30*	-0.002	-0.007	0.000	-0.001	0.329	0.253	0.041	0.039
100	*30*	0.002	0.000	0.001	0.000	0.322	0.246	0.022	0.023
200	*30*	-0.005	-0.005	-0.001	0.001	0.315	0.241	0.016	0.016
30	*50*	-0.001	-0.003	0.001	-0.001	0.244	0.187	0.028	0.028
100	*50*	0.007	-0.001	0.000	0.000	0.241	0.189	0.015	0.015
200	*50*	-0.002	0.004	-0.001	0.000	0.248	0.182	0.010	0.011
30	*100*	0.004	-0.003	0.000	0.000	0.177	0.137	0.018	0.018
100	*100*	0.007	-0.001	0.000	0.000	0.175	0.130	0.010	0.010
200	*100*	0.004	0.000	0.000	0.000	0.171	0.129	0.007	0.007
30	*200*	0.004	-0.002	0.000	0.000	0.121	0.097	0.012	0.012
100	*200*	0.001	-0.002	0.000	0.000	0.122	0.094	0.007	0.007
200	*200*	0.004	0.002	0.000	0.000	0.118	0.092	0.005	0.005

We next compute the bias and the Root mean square error (Rmse) from Monte Carlo repetitions of the Swamy estimates of $\alpha_{ij}, i = 1, \ldots, I, j = 1, 2$ and $\beta_{ij}, i = 1, \ldots, I, j = 1, 2$. The results reported in the table below show that the bias and Rmse of the $\beta_i j$ are fairly small, while the Rmse of the α_{ij} estimates are larger.

Table 1: **Average Adjusted R^2 by rating, sector and maturity bucket**

	Specification		
	1st	2nd	3rd
		Rating	
AAA	0.242	0.190	0.266
AA	0.209	0.366	0.478
A	0.161	0.378	0.463
BBB	0.171	0.307	0.414
		Industrial Sector	
Financials	0.172	0.394	0.505
Industrials	0.177	0.338	0.435
Utilities	0.182	0.295	0.35
		Maturity Bucket	
Short (1-4 years)	0.193	0.402	0.487
Medium (4-10 years)	0.183	0.323	0.438
Long (+10 years)	0.142	0.352	0.44
Overall	0.176	0.354	0.451

Table 1 reports the average adjusted R^2 by rating, sector and maturity bucket for the three specifications and overall.

Table 2: **Principal Components Analysis**

	Panel A	
	Cumulative % explained by PC	
	1st Component	1st and 2nd Component
First Specification (eq. 7)	64.9	83.2
Second Specification (eq. 8)	56.4	82.0
Third Specification	53.8	73.7
	Panel B	
Specification A	63.8	79.1
Specification B	45.6	73.0

Panel A reports the cumulative percentage of the total variation of the residuals, grouped for maturity buckets and industrial sectors, explained by the first two Principal Components (see section 5). By the third specification we mean the specification in eq.(8) augmented by the inclusion of IBOXX triple B index.

.

Table 3: **Cross-section dependence**

Panel A

% PC_1 for y	44.53%
% PC_2 for y	17.92%

Panel B

	average correlation
Delta credit spread	0.40
OLS residuals	0.36
Fixed Effects residuals	0.12
CD stat	55.40 (0.00)
Random Coefficient model (specification B in table 10)	0.07

Table 3 reports the proportion of delta credit spread variability explained by the j-th principal component, PC_j, the average cross-section correlation for delta credit spread and fitted residuals. The statistics in Panel B refers to the residuals of the third specification (eq.(8) plus IBOXX index). CD Stat is the Pesaran (2004) Cross-Section Dependence Statistic. P-value appears in parenthesis.

Table 4: **Description and Predicted effects of the explanatory variables**

Variable	Description	Predicted sign	Remarks
	Individual specific regressors		
cs	Beginning of month credit spread	-	
Avgret	Average of daily excess return over preceding 180 days	-	
Stdret	Standard Deviation of daily excess return over preceding 180 days	+	
rat	Rating	+	We assign a value to each rating. From 10 (AAA) to 1 (BBB-)
Dcsrat	Delta credit spread for rating	+	
Dcsect	Delta credit spread for sector	+	
	Common factors		
5dss	5-year delta swap spread	+	
Nofissue	monthly variation in number of issues included in the IBOXX index	-	
10Gov	10 year German government benchmark monthly variation (mv)	-	
Slope	German government curve slope mv	-	
Conv	German government curve convexity mv		
Upg	monthly variation in upgraded Euro corporate bonds	-	
Downg	monthly variation in downgraded Euro corporate bonds	+	
Mseuro	Morgan Stanley EURO monthly return	-	

Table 5: **Price Estimation Error based on Extended Nelson-Siegel method**

	LEHMAN BROTHERS	SOCIETE TEL FRANCAIS
Coupon	5.47	4.375
Settlement date	31-Aug-05	31-Aug-05
Redemption date	31-Jul-07	12-Nov-10
Redemption value	100	100
Rating	A	A
Observed redemption yield	3.11%	2.99%
Observed price	104.506	106.322
Estimated price	104.190	105.565
Price error	32	76

We estimate the price error for two corporate bonds included in the IBOXX index. The price error is expressed in cents for 100 euros.

Table 6: **Summary Statistics**

		Mean	Std Dev	Min	Max
Credit spread change	*our sample*	-1.58	22.61	-492.20	465.70
	full sample	0.26	44.02	-740.20	2529.80
Coupon (%)	*our sample*	5.55	0.74	3.50	7.25
	full sample	5.37	0.92	2.13	9.75
Years to maturity	*our sample*	5.21	2.28	0.94	14.07
	full sample	5.66	3.45	0.92	29.94
Equity Volatility		1.98%	0.62%	0.00%	6.59%
Equity Excess Return		-0.10%	0.36%	-1.95%	1.27%

Table 6 reports summary statistics on the corporate bonds both in our sample and for all the nonputable and noncallable corporate bonds included in the IBOXX index. The credit spread changes are measured in basis points.

Table 7: **Sample composition for rating and sector**

Panel A		
Rating	% of our sample	% of full sample
AAA	1.80%	5.42%
AA+	0.37%	1.39%
AA	3.35%	5.29%
AA-	13.37%	12.46%
A+	16.35%	15.25%
A	12.88%	14.99%
A-	19.81%	15.00%
BBB+	14.58%	12.67%
BBB	12.87%	12.02%
BBB-	4.63%	5.50%
Panel B		
Industrial Sector	% of sample	% of full sample
Financials	38.16%	37.92%
Industrials	49.76%	50.80%
Utilities	12.08%	11.28%

Table 7 reports summary statistics on the corporate bonds both in our sample and for all the nonputable and noncallable corporate bonds included in the IBOXX index.

Table 8: **Summary Statistics**

Panel A		
Industry	% of sample	% of full sample
Automobiles	9.18%	10.60%
Banks	27.29%	23.66%
Basic-Resources	0.97%	1.31%
Chemicals	1.73%	2.15%
Construction	1.45%	2.38%
Cyclical-Goods & Services	0.97%	1.18%
Energy	2.42%	2.84%
Financial-Services	7.98%	11.45%
Food & Beverage	1.93%	2.19%
Health-Care	1.45%	0.89%
Industrial-Goods & Services	6.73%	6.95%
Insurance	2.90%	2.81%
Media	2.42%	2.10%
Retail	6.84%	6.77%
Telecommunications	13.53%	11.19%
Travel & Leisure	0.16%	0.25%
Utilities	12.08%	11.28%
Panel B		
Maturity Bucket	% of sample	% of full sample
Short (1-4 years)	35.53%	34.84%
Medium (4-10 years)	62.08%	60.06%
Long (+10 years)	2.39%	5.10%

Table 8 reports summary statistics on the corporate bonds both in our sample and for all the nonputable and the noncallable corporate bonds included in the IBOXX index.

Table 9: **Information criteria for common factors**

# of factors	IC_{p1}	IC_{p2}	IC_{p3}
1	8.06	8.06	8.04
2	8.12	8.13	8.10
3	8.23	8.24	8.19
4	8.40	8.42	8.35
5	8.27	8.30	8.21
6	8.34	8.37	8.27
7	8.51	8.54	8.42
8	8.59	8.63	8.50

Table 9 reports the information criteria of Bai and Ng (2002) for detecting the number of common factors in a factor model.

Table 10: **Regression Results**

	A	B	C	D
	Means of estimated coefficients			
cons	-16.381	40.899	-9.757	20.122
	(-1.355)	(1.424)	(-0.880)	(1.030)
cs	-0.185	-0.407	-0.155	-0.256
	(-12.382)	(-20.327)	(-12.090)	(-17.650)
avgret	-925.144	863.641	-968.940	-895.084
	(-1.837)	(0.838)	(-2.130)	(-1.920)
stdret	1228.086	-3285.071	819.269	-1049.308
	(2.035)	(-1.217)	(1.580)	(-0.700)
dcsrat	0.742	0.641	0.708	0.700
	(7.151)	(9.144)	(8.120)	(8.500)
5dss	-3.522	1.171		
	(-0.833)	(0.246)		
nofissue	-46.669	-11.627		
	(-1.842)	(-0.354)		
iret2	11.982	57.065		
	(0.326)	(1.541)		
10gov	4.347	15.696	2.442	2.441
	(1.685)	(4.412)	(1.090)	(1.260)
slope	-6.389	-6.793	-5.000	-8.419
	(-1.762)	(-1.712)	(-1.590)	(-2.630)
conv	-18.061	-21.375	-14.758	-20.829
	(-2.094)	(-2.379)	(-1.370)	(-2.150)
upg	0.226	-2.069	0.127	-1.898
	(0.775)	(-3.607)	(0.460)	(-3.390)
downg	0.438	3.781	0.737	2.694
	(2.400)	(4.695)	(2.590)	(4.090)
mseuro	-15.746	-38.183	-18.993	-30.321
	(-1.678)	(-4.484)	(-2.060)	(-3.670)
factor		-6.927		-4.794
		(-5.204)		(-4.340)
Wald test	417.32	718.86	400.26	691.98
	(0)	(0)	(0)	(0)
Test parameter constancy	5858.91	6084.01	5694.35	7368.95
	(0)	(0)	(0)	(0)
Pesaran-Yamagata Test	8.47	10.30	9.82	11.48
	(0)	(0)	(0)	(0)
Observations	6831	6831	6831	6831

Table 10 reports the estimation results for four specifications with random coefficients. The second column of each specification includes the estimated common factor. t statistics for parameter estimates and p-value for tests appear in parentheses.

Table 11: **Average Partial Correlation of delta credit spreads with the estimated factor**

Rating	average partial correlation
AAA	0.30
AA	0.36
A	0.40
BBB	0.49

Table 11 reports the average correlation of delta credit spreads with the estimated factor, controlling for the explanatory variables contained in Table 10 specification A.

Figure 1: Euro Government Bonds. Market prices and estimated prices

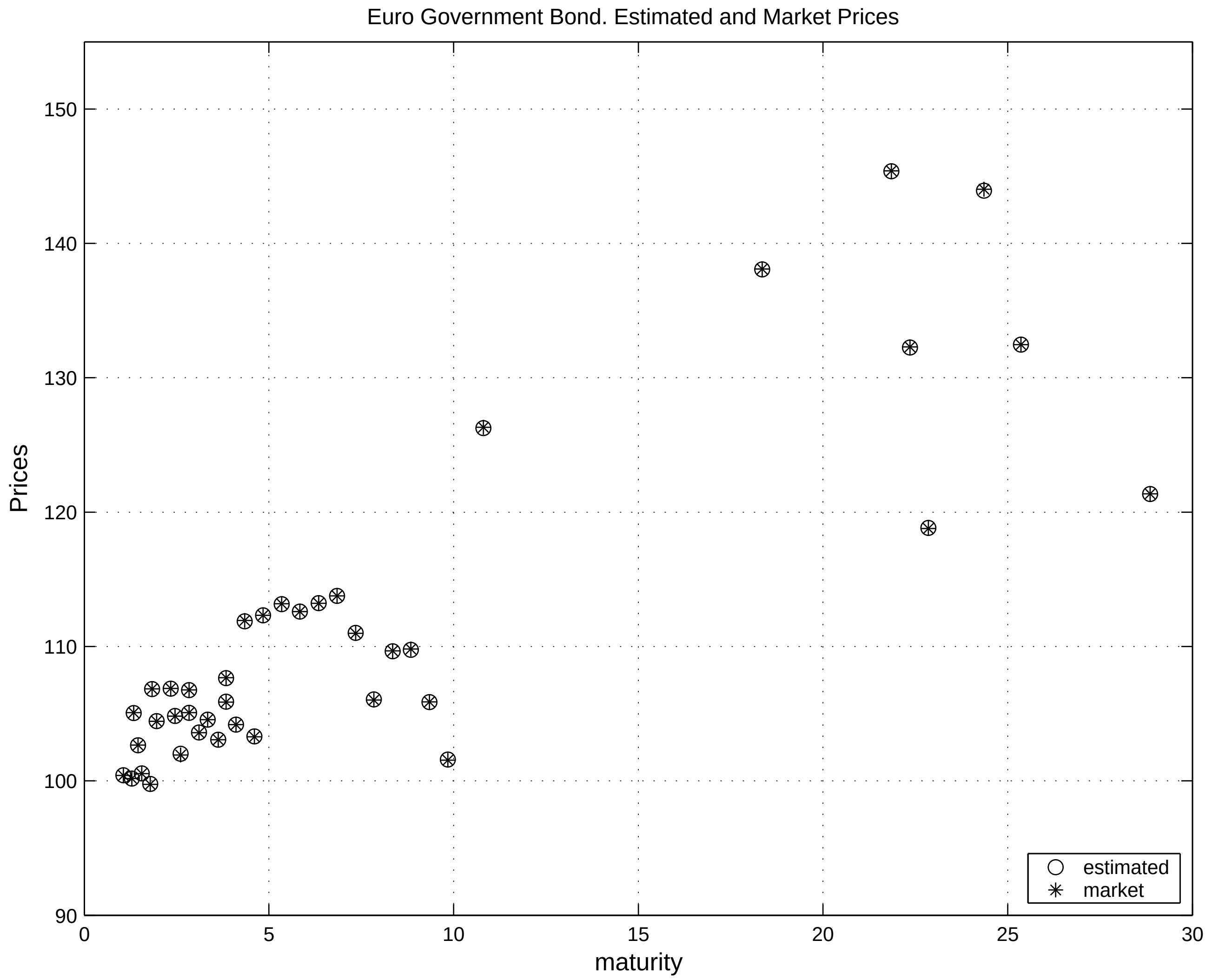

Figure 2: Euro Corporate Bonds. Market prices and estimated prices for A rated bonds