Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Banking & Finance 29 (2005) 1331–1358

Journal of BANKING & FINANCE

www.elsevier.com/locate/econbase

Comparing possible proxies of corporate bond liquidity

Patrick Houweling [a,*], Albert Mentink [b,c,1], Ton Vorst [b,d,2]

[a] *Robeco Asset Management, P.O. Box 973, 3000 AZ Rotterdam, The Netherlands*
[b] *Erasmus University Rotterdam, Rotterdam, The Netherlands*
[c] *AEGON Asset Management, The Netherlands*
[d] *ABN-Amro, The Netherlands*

Received 13 March 2003; accepted 19 April 2004
Available online 3 August 2004

Abstract

We consider nine different proxies (issued amount, listed, euro, on-the-run, age, missing prices, yield volatility, number of contributors and yield dispersion) to measure corporate bond liquidity and use a four-variable model to control for interest rate risk, credit risk, maturity and rating differences between bonds. The null hypothesis that liquidity risk is not priced in our data set of euro corporate bonds is rejected for eight out of nine liquidity proxies. We find significant liquidity premia, ranging from 13 to 23 basis points. A comparison test between liquidity proxies shows limited differences between the proxies.

JEL classifications: C13; G12
Keywords: Liquidity; Corporate bonds; Fama–French model; Euro market

* Corresponding author. Tel.: +31 10 244 3538.
E-mail addresses: p.houweling@robeco.nl (P. Houweling), amentink@aegon.nl (A. Mentink), vorst@few.eur.nl (T. Vorst).
[1] Tel.: +31 70 344 8177.
[2] Tel.: +31 10 408 1270.

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2004.04.007

1. Introduction

The effect of liquidity on bond yields has been frequently studied in the recent finance literature. Since liquidity is a rather subjective concept, a lot of measures have been proposed to approximate the extent to which a bond is liquid or illiquid. For corporate bonds, where most transactions occur on the over-the-counter market, direct liquidity measures (based on transaction data) are often not reliable and difficult to obtain. Therefore, researchers resorted to indirect measures ('proxies') that are based on bond characteristics and/or end-of-day prices. This paper makes a number of contributions to this literature on measuring corporate bond liquidity. First, we pay great attention to control for other sources of risk than liquidity to properly identify the premium that is associated with liquidity risk. As far as we know, this is the first study in this strand of the literature to use the well-known Fama and French (1993) two-factor bond-market model to control for interest rate and credit risk and to augment it with individual bond characteristics, rating and maturity, as recommended by Gebhardt et al. (2001). Second, we do not make a subjective choice of which liquidity proxies to work with, but implement as much of the proxies proposed in the literature as possible on our data set. We evaluate the relative performance of all proxies, employing a method recently applied by Goldreich et al. (2002) on Treasury bonds. Third, the vast majority of empirical papers on sovereign and corporate bond liquidity studied data from the United States and relatively little is known about the extent to which these results apply to the euro market. Although euro corporate bond data were also studied by other authors, including Annaert and De Ceuster (1999), McGinty (2001) and Díaz and Navarro (2002), none of them analyzed the euro corporate bond market using data on individual bonds over a substantial time period.

We use the Brennan and Subrahmanyam (1996) methodology of liquidity-sorted portfolios to test whether liquidity is priced in the euro-denominated corporate bond market. We use nine proxies for bond liquidity: *issued amount*, *listed*, *on-the-run*, *euro*, *age*, *missing prices*, *yield volatility*, *number of contributors* and *yield dispersion*; see Section 3.4 for a detailed description. For each liquidity proxy, we construct P, mutually exclusive portfolios by sorting all bonds on their value of the liquidity proxy and assigning the first $100/P\%$ of the bonds to portfolio 1, the next $100/P\%$ to portfolio 2, and so on, until the last $100/P\%$ of the bonds are assigned to portfolio P. The P time series of portfolio yields are subsequently used in two regression models. In the first model, each portfolio has a constant liquidity premium. In the second model, the liquidity premium is time-varying and a function of the size of liquidity proxy. In both models, the null hypothesis states that the portfolios' liquidity premiums are jointly equal to zero. We use a detailed data set consisting of daily yields of individual corporate bonds which are denominated in euros or in one of the currencies of the euro-participating countries ('legacy' currencies). The results for the first regression model indicate that the null hypothesis of no liquidity premium is rejected for eight out of nine liquidity proxies. So, we find strong evidence of priced liquidity. For the second model, the null hypothesis of no liquidity effects is even always rejected. To determine the relative effectiveness of the different liquidity proxies, we run a series of regressions with pairwise combinations of the liquidity proxies, as proposed by Goldreich et al. (2002). This allows us to rank the different liquidity proxies we consider. The results of the tests point out that no proxy stands out from the rest.

The remainder of this paper is structured as follows. Section 2 gives an overview of the methodologies and results of the empirical liquidity literature. Section 3 describes how we control for other sources of risk than liquidity risk and how we estimate the liquidity premium. This section also describes the portfolio construction and our nine liquidity proxies. Next, Section 4 describes the data that are used to test the hypotheses of corporate bond liquidity. Section 5 presents the results from the model implementation. Finally, Section 6 summarizes the paper.

2. Literature

Both theoretical and empirical evidence demonstrate that liquidity risk is priced in security markets. The market microstructure models of Amihud and Mendelson (1986), Boudoukh and Whitelaw (1993) and Vayanos (1998) show that transaction costs cause liquidity differences between securities, and that illiquid securities have higher expected rates of return than liquid securities.

For equity markets, empirical evidence on priced liquidity risk is provided by, e.g., Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Haugen and Baker (1996), Brennan et al. (1998), Chordia et al. (2001) and Chordia et al. (2001). For bond markets, a substantial part of empirical studies analyzed data from the US Treasury market, where bonds are issued on a regular basis and price data are easily available. Also, controlling for other sources of risk than liquidity risk is relatively easy in this market, because credit risk is not an issue. To control for interest rate risk, authors have used several approaches. The first approach is to create pairs of zero-coupon bonds with exactly the same maturity date; this fully eliminates interest rate risk. Amihud and Mendelson (1991), Kamara (1994) and Strebulaev (2001) used this method to test for liquidity differences between US Treasury notes and bills; Fleming (2002) compared US Treasury bonds with small and large outstanding amounts. The second approach is to form triplets of coupon bonds, which, with suitable bond weights, also eliminates interest rate risk. Elton and Green (1998) used this method to examine yield differences between bonds with high and low trading volume. Another frequently used approach is to analyze the yield difference between the *on-the-run* (most recently issued) bond and *off-the-run* (older) bonds; this will, however, leave a small maturity gap between the bonds. Warga (1992), Goldreich et al. (2002) and Krishnamurthy (2002) used this method on US Treasury data and Boudoukh and Whitelaw (1991, 1993) on Japanese data. All papers mentioned above, except Strebulaev (2001), found statistically significant liquidity premiums.

Research on corporate bond liquidity is substantially more difficult, because of the presence of credit risk and the smaller number of bonds per issuer. A strategy of matching bonds by maturity and issuer, similar to the Treasury studies above, will typically generate too few observations. As far as we know, there is only one study that successfully applied this approach: Crabbe and Turner (1995) analyzed pairs of new issues, issued by the same borrower, with identical issue and maturity dates, but with different issue sizes. The most popular approach is to regress yields (and occasionally bid–ask spreads or trading volumes) of individual corporate bonds on a range of proxies for interest rate, credit and liquidity risk. Examples of studies that used this method include Gehr and Martell (1992), Shulman et al. (1993), Chakravarty and Sarkar (1999), Alexander et al. (2000), Hong and Warga (2000), Collin-Dufresne et al. (2001), Ericsson and Renault (2001), Schultz (2001), Díaz and Navarro (2002), Elton et al. (2002) and Mullineaux and Roten (2002). Remarkably, all papers studied US data, except Díaz and Navarro (2002) who studied Spanish corporate bonds. Cornell (1992), Fridson and Jonsson (1995) and Annaert and De Ceuster (1999) used similar regression approaches, but on *indices* of US mutual funds, US high yield bonds and euro investment grade bonds, respectively. Finally, McGinty (2001) analyzed one month of euro corporate bond data using scatter plots and tables. All papers mentioned above, except for Gehr and Martell (1992) and Crabbe and Turner (1995), found evidence of significant liquidity premiums for at least one liquidity proxy.

To summarize, almost all empirical papers on bond liquidity found significant liquidity effects for government and corporate bonds. However, none of the studies used the portfolio-based testing methodology often employed in the literature on equity liquidity. Moreover, although there is ample research on the US market, the evidence for euro-denominated bonds is limited to papers that study index data (Annaert and De Ceuster, 1999), a small sample period (McGinty, 2001) or data from one country (Díaz and Navarro, 2002).

3. Methodology

This section describes the methodology used to test whether liquidity risk is priced in the euro-denominated corporate bond market. First, we explain how we control for other sources of risk than liquidity risk using Fama and French (1993) and Gebhardt et al. (2001). Next, we describe the implementation of our models and the Goldreich et al. (2002) method to compare different liquidity proxies. Finally, we present our liquidity proxies.

3.1. Controlling for other sources of risk

In measuring a security's liquidity premium, it is important to realize that the security's expected return is not only affected by liquidity risk but also by other sources of risk. Theory (like the reduced form credit risk models following Jarrow and Turnbull, 1995) nominates two risk factors: (i) interest rate risk and (ii) credit risk. We use the Fama and French (1993) bond-market model as a starting point to proxy for interest rate and credit risk. They found two risk factors that explained over 90% of the variation in realized *excess returns* on corporate bond portfolios; the excess return was defined as the portfolio return minus the one-month Treasury rate. The first risk factor was calculated as the long-term Treasury bond return minus the one-month Treasury rate at the end of the previous period. Thus, this *slope factor* should explain variations in excess bond returns by changes in the slope of the Treasury yield curve. The second factor was defined as the return on a market portfolio of long-term corporate bonds minus the long-term Treasury bond return. This *credit factor* was therefore related to the likelihood of credit events in the corporate bond portfolio.

Unlike Fama and French (1993), we do not use a bond's realized return as proxy for its expected return, but, following the bond liquidity literature, we use the bond's yield-to-maturity. [3] The advantage of yields is that they are forward-looking, while realized returns are backward looking. In all regressions, we thus replace the excess realized return by the *excess yield*, which is defined as the yield minus the short-term default-free rate.

A second modification to the Fama–French model concerns the choice of the default-free interest rate curve, which is required to calculate the excess yields and the two risk factors. Instead of using the government curve, we use the swap curve. Our motivation is that since the end of the 1990s, fixed-income investors have moved away from using government securities to extract default-free interest rates and started using interest rate swap rates instead; see also Golub and Tilman (2000) and Kocić et al. (2000). In Section 5.1, we test both proxies for default-free rates.

Gebhardt et al. (2001) looked at the validity of the Fama–French bond-market model by analyzing whether individual bond characteristics could rival the two Fama–French factors. Three characteristics were considered: rating, duration and Altman (1968) *Z*-scores. They concluded that both Fama–French factors and bond characteristics were important in explaining bond yields and recommended a model containing four variables: the Fama–French slope and credit factors, rating and duration. In Section 5.2, we show that for our data set these four characteristics are also relevant. [4] Therefore, our null model to control for other sources of risk consists of four variables: two Fama–French factors and two characteristics; the model is described formally in Section 3.2. Clearly, all our conclusions about the relation between liquidity and bond yields are based on the assumption that our four-variable pricing model correctly and fully controls for interest rate and credit risk; see also Dimson and Hanke (2001). [5]

To the best of our knowledge, no other paper in the liquidity literature has employed both the Fama–French factors and individual bond characteristics to control for other sources of risk. One paper, Ericsson and Renault (2001), used the Fama–French factors, but not the characteristics; several papers made use of the rating and

[3] Changing the return measure from realized return into yield makes our return measure sensitive to time-to-maturity. Therefore, we incorporate time-to-maturity in our regression models as control variable; see also the discussion of the Gebhardt et al. (2001) paper below.

[4] We have replaced duration by maturity, but this should not affect our results as both variables are highly correlated.

[5] We further assume that taxes do not affect bond yields.

maturity characteristics, including Alexander et al. (2000), Hong and Warga (2000) and Mullineaux and Roten (2002), but not of the Fama–French factors; most papers used a list of ad-hoc proxies.

3.2. Models

Unlike prior papers on bond liquidity, we do not estimate our models on individual bonds, but on constructed portfolios, like in the equity literature; see e.g. Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996) and Haugen and Baker (1996). Specifically, we follow Brennan and Subrahmanyam (1996) by creating liquidity-sorted portfolios and testing whether the constructed portfolios have significantly different yields, while controlling for other sources of risk as described above. From the literature, we collect nine liquidity proxies, which are detailed in Section 3.4. For each proxy i, we create P mutually exclusive portfolios as follows (the choice for P will be discussed at the end of this section).

Every two weeks, we order all bonds in the sample by their value of liquidity proxy i; only bonds that have already been issued and have not yet matured on that date are used in the ordering. Then, we assign the first $100/P$% of the bonds to portfolio 1, the next $100/P$% to portfolio 2, and so on, until the last $100/P$% of the bonds are assigned to portfolio P. The sort order is chosen such that portfolio 1 contains the bonds that proxy i hypothesizes to be the most liquid and portfolio P the most illiquid. Every day we calculate the yield of each portfolio as the unweighted average of the yields of the bonds that make up the portfolio. A bond's yield is determined as follows: if the bond is not quoted, we disregard it for that day; if it is quoted by one pricing source, we use that yield; if it is quoted by more than one pricing source, we use the average quote. For each proxy i, we have now created P time series of portfolio yields.

As in Brennan and Subrahmanyam (1996), the time series are used in two regression models. In the first model, each portfolio has a constant liquidity premium. Formally, model 1 is as follows: [6]

$$\begin{aligned} &Y^i_{pt} = \alpha^i_p + \sum_{j=1}^{2} \beta^i_{jp} F_{jt} + \sum_{j=1}^{2} \gamma^i_j C^i_{jpt} + \varepsilon^i_{pt}, \\ &E[\varepsilon^i_{pt}] = 0, \\ &E[\varepsilon^i_{pt}\varepsilon^i_{qs}] = \sigma^i_{pq}, \text{ if } t = s, \text{ and 0 otherwise,} \end{aligned} \tag{1}$$

where superscripts i refer to liquidity proxy i, Y^i_{pt} is the excess yield of the pth proxy-i portfolio on day t, F_{1t} and F_{2t} are the two Fama–French factors and C^i_{1pt} and C^i_{2pt} are the two portfolio characteristics. The coefficients are interpreted as follows: α^i_p is the portfolio-specific liquidity premium, β^i_{jp} is the portfolio-specific factor loading for

[6] As suggested by an anonymous referee, we have extended models 1 and 2 with quadratic terms of the two portfolio characteristics, rating and maturity. The results showed to be robust with respect to this extension of both models.

Fama–French factor j and γ^i_j is the marginal effect of portfolio characteristic j. The model controls the portfolio excess yields for the two Fama–French factors and the two bond characteristics. Thus, it gives the effect of a particular liquidity proxy after correcting for the four features. Note that the Fama–French factors have portfolio-specific coefficients and common variable values, while the characteristics have common coefficients and portfolio-specific variable values.

The disturbance terms are allowed to be heteroscedastically distributed and cross-sectionally correlated, but we do assume that they are uncorrelated across time. To correct for possible autocorrelations in the disturbances, we apply the Newey and West (1987) estimator for the covariance matrix. For proxy i, we estimate all $3P+2$ coefficients $(\alpha^i_1,\ldots,\alpha^i_P,\ \beta^i_{11},\ldots,\beta^i_{1P},\beta^i_{21},\ldots,\beta^i_{2P},\ \gamma^i_1,\gamma^i_2)$ for all P portfolios *simultaneously* with Feasible Generalized Least Squares (FGLS) as a system of seemingly unrelated regressions (SUR); see e.g. Greene (2000, Chapter 15).

To test the null hypothesis that proxy i has no liquidity premium, or in other words, that the two Fama–French factors and the two portfolio characteristics fully explain the bond yields, we use a Wald test to determine the joint significance of the intercepts: $H_0 : \alpha^i_1 = 0 \wedge \ldots \wedge \alpha^i_P = 0$. The test statistic is asymptotically χ^2-distributed with P degrees of freedom. [7]

In the second model, we change the functional form of the liquidity premium: all portfolios share a common intercept and a portfolio-specific liquidity variable is added to the regression equation. Formally, regression model 2 reads

$$Y^i_{pt} = \alpha^i + \sum_{j=1}^{2} \beta^i_{jp} F_{jt} + \sum_{j=1}^{2} \gamma^i_j C^i_{jpt} + \delta^i L^i_{pt} + \varepsilon^i_{pt}, \tag{2}$$

where the definitions of the Fama–French factors, the portfolio characteristics and the assumptions on the disturbances are equal to those in Eq. (1) and L^i_{pt} is the value of the liquidity proxy for the pth proxy-i portfolio on day t in deviation from its daily average; so, if l^i_{pt} denotes the value of the liquidity proxy, and $\bar{l}^i_t$ is its daily average, i.e.

$$\bar{l}^i_t = \frac{1}{P} \sum_{p=1}^{P} l^i_{pt},$$

then L^i_{pt} is calculated as $L^i_{pt} = l^i_{pt} - \bar{l}^i_t$. We have chosen this normalization of the liquidity proxy to correct for a possible change in the mean during our sample period. For example, the average amount outstanding has risen from €353 million on the first day of our sample to €434 million on the last day.

[7] There is caveat in the interpretation of the test results: if we want to test whether proxy i is a good liquidity proxy, we are actually testing a joint hypothesis: illiquidity leads to yield increases *and* proxy i is a proxy for liquidity; see also Kempf and Uhrig-Homburg (2000) and Jankowitsch et al. (2002). If we reject this joint hypothesis, then either illiquidity does not lead to yield increases or i is not a good liquidity proxy (or both). Given the strong empirical evidence mentioned in Section 2, we feel confident that a rejection of the joint hypothesis can in fact be traced to i being an inadequate liquidity proxy.

In Eq. (2), the portfolio-specific intercepts of Eq. (1) have been replaced by a single intercept and an additional regressor has been introduced that contains a proxy for portfolio p's liquidity. This changes the functional form of the liquidity premium: the constant liquidity premium of α_p^i in model 1 has been replaced by a time-varying premium $\alpha^i + \delta^i L_{pt}^i$ that is linear in the value of the liquidity proxy (in deviation from its mean). Here, the null hypothesis of no liquidity effect is tested with a Wald test on the joint significance of α^i and δ^i: H_0: $\alpha^i = 0 \wedge \delta^i = 0$. The test statistic is asymptotically χ^2-distributed with 2 degrees of freedom. The joint hypothesis problem discussed in footnote 7 also applies here.

We now discuss the choice for the number of portfolios P for both models. For model 1, we create two portfolios for each liquidity proxy. This gives an intuitive interpretation of portfolio 1 as the 'liquid portfolio' and portfolio 2 as the 'illiquid portfolio'. Moreover, the difference $\alpha_2^i - \alpha_1^i$ between the two intercepts can be interpreted as the yield premium investors get for bearing liquidity risk caused by proxy i. In model 2, we have to estimate the slope coefficient δ^i, i.e. the relation between a portfolio's value for liquidity proxy i and its excess yield. Clearly, two portfolios would be insufficient to estimate a slope. However, using 'too much' portfolios diminishes the power of the Wald test; see Lys and Sabino (1992). From their Fig. 1, it follows that if the portfolios contain approximately 25% of the bonds, the power of the test of no relation between the liquidity proxy and the excess yield is maximized. Therefore, we use four portfolios for model 2.

3.3. Comparison

Given the large number of liquidity proxies that has been proposed in the literature, a natural question to ask is whether all proxies are equally suited to proxy bond liquidity or if some proxies work better than others. We follow Goldreich et al. (2002) by running a series of regressions with pairwise combinations of the liquidity proxies. For each combination (i, k) of proxies, we estimate a regression like Eq. (2) for proxy i, augmented with proxy k:

$$Y_{pt}^i = \alpha^i + \sum_{j=1}^{2} \beta_{jp}^i F_{jt} + \sum_{j=1}^{2} \gamma_j^i C_{jpt}^i + \delta^i L_{pt}^i + \delta^{ik} L_{pt}^{ik} + \varepsilon_{pt}^{ik}, \tag{3}$$

where L_{pt}^{ik} is the value of liquidity proxy k for the pth proxy-i portfolio in deviation of its daily average. Further, the coefficients are defined and disturbances behave as in Eq. (1).

In this regression equation, [8] we test for the significance of δ^{ik}. If it is significant, we say that 'k adds explanatory power to i', and otherwise we say that 'k is subsumed by i' (this follows the terminology in Goldreich et al. (2002)). By repeating this procedure for all possible combinations, we can count the number of times a proxy adds

[8] Goldreich et al. (2002) first orthogonalized the values of proxy k relative to proxy i and used the orthogonalized values in Eq. (3) instead of L_{pt}^{ik}. This is not necessary, since, by the Frisch–Waugh theorem (see e.g. Greene, 2000, Section 6.4.3), the regression already 'automatically' does this for us.

power to another proxy, and the number of times a proxy subsumes another proxy. This allows us to rank the different liquidity proxies we consider.

3.4. Liquidity proxies

Empirical papers that examined liquidity in bond or equity markets used both *direct* measures (based on transaction data) and *indirect* measures (based on bond characteristics and/or end-of-day prices). Examples of direct liquidity measures are *quoted bid–ask spreads*, *effective bid–ask spreads*, *quote sizes*, *trade sizes*, *quote frequencies*, *trade frequencies* and *trading volume*. For corporate bonds, where most transactions occur on the over-the-counter market, these direct measures are often not reliable and difficult to obtain. Therefore, we use indirect liquidity proxies instead. By searching the theoretical and empirical liquidity literature, we found nine liquidity proxies that can be implemented on our data set. [9] Table 1 shows which papers used which proxies and the effects they found; three proxies, *euro*, *missing prices* and *yield dispersion*, are not mentioned in this table, because they were not used in previous studies. We will now discuss each proxy in more detail, elaborating on their interpretation, their expected effect on bond yields and theoretical and/or empirical evidence.

3.4.1. Issued amount

The *issued amount* of a bond is often assumed to give an indication of its liquidity. Most investment banks use it as liquidity criterion in building their bond indices; for example, Lehman Brothers uses this criterion for their Euro-Aggregate Corporate Bond index. Its use was first proposed by Fisher (1959), who claimed that large issues should trade more often, so that the proxy *issued amount* is actually a proxy for the direct liquidity measure trading volume. Since Fisher, several alternative hypotheses have been put forward that also predict a positive effect of *issued amount* on liquidity (and thus on bond prices). In market microstructure models, like Smidt (1971) and Garman (1976), transaction costs arise, because dealers hold inventories. Further, dealers' inventory costs are higher if it is more difficult to obtain information about a security and if the expected holding time is longer. Crabbe and Turner (1995) subsequently reasoned that large issues may have lower information costs, since more investors own them or have analyzed its features; similarly, information about small issues may be less broadly disseminated among investors. Therefore, small issues will have a higher yield due to an illiquidity premium. Another frequently heard argument, for instance in Sarig and Warga (1989) and Amihud and Mendelson (1991), is that bonds with smaller issued amounts tend to get locked in buy-and-hold portfolios more easily, reducing the tradable amount and thus their liquidity. To summarize the above, we hypothesize a negative effect of *issued amount* on yields.

[9] We would like to stress that by selecting liquidity proxies from theoretical research and from empirical research on other data bases, the effects of *data-snooping* on portfolio-based tests, as described by Lo and MacKinlay (1990), are probably limited.

Table 1
Overview of liquidity proxies from the empirical bond liquidity literature[a]

Authors[b]	Data[c]	Liquidity proxies					
		Issued amount	Listed	On-the-run	Age	Yield volatility	Number of contributors
Corporate bonds							
AEF00	US	$-^{*}$	$+^{*}$		$+^{*}$	$-^{*}$	
CT95	US	$\diamond$					
EGAM02	US	$\diamond$			$+^{*}$		
ER01	US				$+^{*}$		
GM92	US	$+^{\diamond}$ $-^{\diamond}$					$-^{\diamond}$
HW00	US	$-^{*}$			$+^{*}$	$+^{*}$	
M01	EMU	$\diamond$			$\diamond$		
MR02	US	$+^{\diamond}$ $-^{\diamond}$					
S01	US				$+^{\diamond}$		
SBP93	US	$\diamond$				$+^{*}$	
Treasury bonds							
AM91	US						
EG98	US			$+^{*}$	$+^{*}$		
F02	US	$\diamond$		$+^{*}$			
JMP02	EMU[d]	$-^{*}$		$+^{*}$			$-^{*}$
K02	US	$-^{*}$					
KU00[e]	Germany	$-^{*}$					
SW89	US	$-^{\diamond}$			$+^{*}$		
W92	US	$-^{\diamond}$			$+^{*}$		
Corporate and Treasury bonds							
DN02	Spain	$+^{*}$ $-^{*}$			$+^{*}$		
Corporate, municipal and Treasury bonds							
CS99	US				$+^{*}$		

[a] Legend: – negative; + positive; * significant; $\diamond$ insignificant.

[b] AEF00 = Alexander et al. (2000), AM91 = Amihud and Mendelson (1991), CS99 = Chakravarty and Sarkar (1999), CT95 = Crabbe and Turner (1995), DN02 = Díaz and Navarro (2002), EG98 = Elton and Green (1998), EGAM02 = Elton et al. (2002), ER01 = Ericsson and Renault (2001), F02 = Fleming (2002), GM92 = Gehr and Martell (1992), HW00 = Hong and Warga (2000), JMP02 = Jankowitsch et al. (2002), K02 = Krishnamurthy (2002), KU00 = Kempf and Uhrig-Homburg (2000), M01 = McGinty (2001), MR02 = Mullineaux and Roten (2002), S01 = Schultz (2001), SBP93 = Shulman et al. (1993), SW89 = Sarig and Warga (1989), W92 = Warga (1992).

[c] EMU = European Monetary Union, US = United States.

[d] JMP02 considered 6 countries: Austria, France, Germany, Italy, Spain and The Netherlands.

[e] We used the price discounts in KU00's Table 2 to calculate the impact of maturity on yields.

Table 1 shows that many empirical papers considered *issued amount* as liquidity proxy. The papers on Treasury bonds found negative and mostly significant effects, so that larger Treasury issues have lower yields, as expected. Research on corporate bonds is inconclusive, though: both negative and positive coefficients are observed. McGinty (2001) confirmed this by showing that even though most large issues in his corporate bond sample were liquid, some large issues were illiquid and some small issues were liquid.

3.4.2. Listed

Alexander et al. (2000) reasoned that companies whose equity is listed on a stock exchange must disclose more information than privately held companies. According to the market microstructure models mentioned above, the costs of making a market in bonds of listed firms should thus be smaller. Therefore, we hypothesize that the proxy *listed* is associated with higher liquidity and lower yields.

Since Alexander et al. (2000) were the only authors to use the liquidity proxy *listed*, the empirical evidence is limited to their results. Contrary to their expectations, they found that issues of private firms trade more actively and thus are more liquid than issues of listed firms. Their explanation of this result was that for private firms debt is the only investment vehicle, while for public firms both debt and equity are traded; therefore, debt of private firms might trade more and have higher liquidity.

3.4.3. Euro

The next liquidity proxy is whether a corporate bond is denominated in euros or in one of the legacy currencies. The market generally sees legacy bonds (i.e. denominated in one of the currencies of the euro-participating countries) as the less liquid ones, because these bonds are relatively old, not well known to the bond investors and more difficult to trade. The predicted sign of the proxy *euro* is thus higher liquidity and lower yields. This bond characteristic splits the corporate bond sample into two excluding groups: euro bonds and legacy bonds. To our best knowledge, no other papers have implemented this liquidity proxy.

3.4.4. On-the-run [10]

In general, on-the-run, or most recently issued, bonds are considered to be the most liquid bonds, in contrast to off-the-run or older bonds, as market participants often focus their attention on younger bonds. For each issuer, we define the bonds that are issued most recently as on-the-run bonds and the remaining, older bonds as off-the-run bonds. In case an issuer has issued only one bond, we define this bond as an on-the-run bond. [11] We test the hypothesis that on-the-run bonds have a higher liquidity, and consequently a lower yield, than the off-the run bonds.

This liquidity proxy has been implemented in the empirical literature, as Table 1 displays, but only for Treasury bonds. All papers found a positive and statistically

[10] We thank an anonymous referee for this valuable suggestion.

[11] By definition, the distinction between on-the-run and off-the-run bonds is related to *age*. Yet, portfolios that are constructed with the help of each proxy do differ. Using the liquidity proxy *on-the-run*, only bonds possessing this feature, thus both old and young, end up in the liquid portfolio, whereas applying the liquidity proxy *age* the liquid portfolio is only composed of young bonds, both on-the-run and off-the-run.

significant effect of *on-the-run*. As far as we know, no prior papers implemented this proxy on corporate bonds.

3.4.5. Age

The *age* of a bond is a popular proxy for its liquidity. Sarig and Warga (1989) observed that as a bond gets older, an increasing percentage of its issued amount is absorbed in investors' buy-and-hold portfolios. Thus, the older a bond gets, the less trading takes place, and the less liquid it becomes. Moreover, once a bond becomes illiquid, it stays illiquid until it matures. McGinty (2001) and Schultz (2001) also noted that new issues trade more than old issues. McGinty mentioned lead managers' commitment to making market in the newly issued bond. Schultz pointed out that new issues are typically under priced, so that traders buy bonds after the offering and sell them shortly thereafter. Following these arguments, we hypothesize a positive relation between *age* and yield.

Empirical research strongly confirms the positive effect of *age* on yields; see Table 1. This finding holds for corporate and sovereign bonds and for US and European data sets. Moreover, Schultz (2001) found evidence for the argument by Sarig and Warga (1989), since in his sample most bonds were bought and not sold; in other words, the bonds were put in buy-and-hold portfolios.

Market practitioners often use a threshold value to determine if a bond is 'old' or 'young': for some T, they mark all bonds with an age smaller than T as 'young' and an age larger than T as 'old'. Some academic papers also use such a dichotomous approach for the liquidity proxy *age*. For instance, Alexander et al. (2000) set $T=2$ years, Ericsson and Renault (2001) used $T=3$ months, and Elton et al. (2002) employed a threshold value of 1 year. To determine which threshold values give a useful division of bonds, we estimate model 1 from two portfolios, where portfolio 1 contains all bonds younger than T months and portfolio 2 older than T months, for $T=2,4,\ldots,30$. The difference $\alpha_2-\alpha_1$ between the portfolio intercepts, i.e. the liquidity premium between old and young bonds, and the significance of the Wald test on H_0: $\alpha_2-\alpha_1=0$ are displayed in Fig. 1. Thresholds from 4 to 24 months give rise to a significant liquidity premium, while the 2-month threshold and thresholds larger than 24 months do not. The division between young and old bonds seems to be the strongest for a threshold of 14 months, where the premium equals 36 bps. For the remainder of this study, we arbitrarily use a threshold of 1 year for the proxy *age*, although any other value between 4 months and 2 years could also be used.

3.4.6. Missing prices

The occurrence of 'price runs' and missing values is our first liquidity proxy that uses market information. Sarig and Warga (1989) argued that if the liquidity of a bond is sufficiently low, it may happen that on some business days there is virtually no trading in that bond. In their data set, this was recorded as a 'price run': two consecutive prices for a bond were identical. We extend their notion of illiquidity by considering not only the occurrence of a price run, but also the occurrence of a missing value, since in both cases there is no activity in that bond on that day. We will jointly

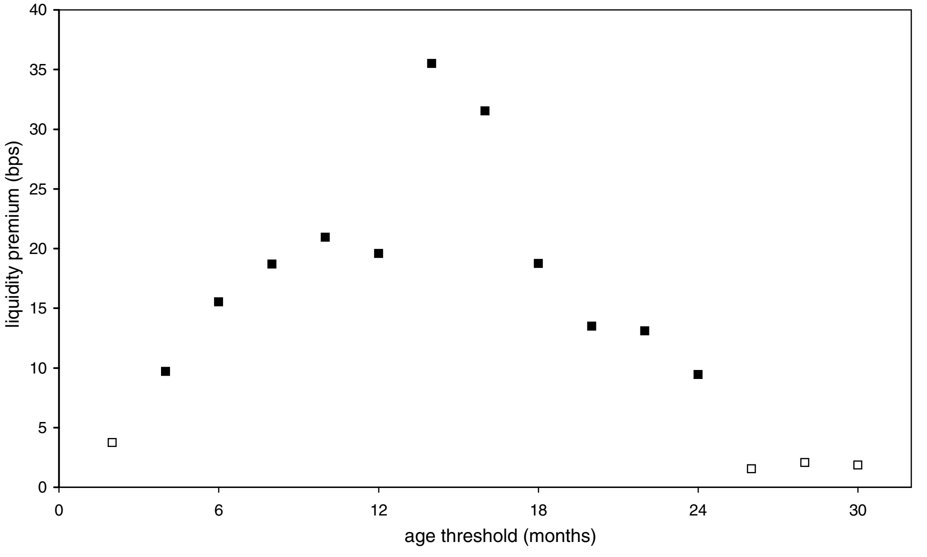

Fig. 1. Liquidity premiums for different age thresholds. Solid (■) and empty (□) squares denote significance and insignificance, respectively, of the Wald test on the joint significance of the two intercepts (p-value < 0.05).

refer to these events as the proxy *missing prices*. We hypothesize a positive relation between *missing prices* and yield.

3.4.7. Yield volatility

The proxy *yield volatility* is a measure of yield uncertainty. In the market microstructure models discussed above, dealers' inventory costs are higher if information uncertainty is higher. An important source of uncertainty is related to the predictability of future yield movements. Therefore, we hypothesize that a higher *yield volatility* leads to larger bid–ask spreads, and thus to lower liquidity and higher yields.

The empirical evidence for yield uncertainty as liquidity proxy is mixed; see Table 1. Shulman et al. (1993) used *price volatility* as proxy for price uncertainty and found a significantly positive effect on bond spreads. Hong and Warga (2000) proxied uncertainty with squared price return and estimated a positive and significant coefficient in a regression using bid–ask spread as dependent variable; this also implies a positive effect of uncertainty on bond yields. Alexander et al. (2000) approximated uncertainty as the average of absolute price returns; in their regressions, they found a significant, positive effect on trading volume, implying a negative relation between uncertainty and yields.

3.4.8. Number of contributors

The *number of contributors* is our following proxy for a bond's liquidity, and the first that uses quote composition information. In Ericsson and Renault (2001), a larger number of active traders competing for the same bond leads to a smaller price discount for illiquidity and thus a smaller yield premium. Alternatively, Gehr and Martell (1992) and Jankowitsch et al. (2002) argued that a larger number of market participants makes it easier to trade a bond, because it is easier to find a counter party for a transaction and large orders can be split up into smaller parts without affecting the market price. Either way, we hypothesize a positive relation between the proxy *number of contributors* and liquidity and therefore expect a negative effect of this proxy on bond yields.

Direct empirical evidence on the *number of contributors* liquidity proxy is limited. Jankowitsch et al. (2002) found that bonds with more contributors have lower yields for all but one of the six European countries they analyzed. Indirect evidence is provided by Schultz (2001), who showed that there was a positive relation between the number of trades in a bond and the number of dealers as counter parties. Further, the results of Gehr and Martell (1992) showed a negative, though insignificant effect of the number of dealers on the bid–ask spread.

3.4.9. Yield dispersion

Our final liquidity proxy, *yield dispersion*, reflects the extent to which market participants agree on the value of a bond. Tychon and Vannetelbosch (2002) derived a model that predicts that if investors have more heterogeneous beliefs, the liquidity premium is larger. The inventory costs argument, mentioned above, applies here as well, since dealers face more uncertainty if prices show a larger diffusion among contributors. Either way, we hypothesize a positive relation between dispersion and bond yields.

We proxy this notion of liquidity with a *yield dispersion* statistic, which has not been used before in the literature, as far as we know. We define the yield dispersion of bond b on day t as the standard deviation of percentage yield differences relative to the mean:

$$\text{Dispersion}_{bt} = \sqrt{\frac{1}{n_{bt}-1}\sum_{s=1}^{n_{bt}}\left(\frac{y_{bts}-\bar{y}_{bt}}{\bar{y}_{bt}}\right)^2}, \tag{4}$$

where y_{bts} is the yield quoted by pricing source s, $\bar{y}_{bt}$ is the average yield and n_{bt} is the number of contributors. This proxy can only be calculated if we have at least two quotes for a bond on a particular day, i.e. if $n_{bt} > 1$.

3.4.10. Application

Table 2 gives details on the calculation of each liquidity proxy. It also shows the expected sign of the proxy. To get the l^i_{pt} variable of Section 3.2, we multiply proxies with a negative expected sign by -1. After this transformation, the δ^i coefficient of model 2 is hypothesized to be positive for all proxies; this facilitates checking the results with the hypotheses. Finally, the table shows the order in which bonds are put in the portfolios: the first portfolio always contains the bonds that are hypothesized to be most liquid, the last portfolio contains bonds that we expect to be most illiquid.

As described in Section 3.2, every two weeks the portfolios for each liquidity proxy are rebalanced according to each bond's value for that proxy. For the proxies

Table 2
Overview of liquidity proxies, their expected signs and the portfolio order

Liquidity proxy	Details	Sign[a]	Portfolio[b]	
			First	Last
Issued amount	Total notional in billions of euros	–	Largest	Smallest
Listed	1 if a firm's equity is publicly traded, 0 otherwise	–	Yes	No
Euro	1 if bond is denominated in euros, 0 otherwise	–	Euro	Legacy
On-the-run	1 if bond is on-the-run, 0 otherwise	–	On-the-run	Off-the-run
Age	Time between issue date and quote date in years	+	Young	Old
Missing prices	1 if price is missing or equal to previous price, 0 otherwise	+	Least	Most
Yield volatility	Standard deviation of yields since previous rebalancing	+	Lowest	Highest
Number of contributors	Number of market participants quoting the bond	–	Largest	Smallest
Yield dispersion	See Eq. (4)	+	Smallest	Largest

[a] Expected signs of the relationship between the proxies and bond yields.
[b] Order in which the ranked bonds are assigned to the first (most liquid) portfolio and the last (most illiquid) portfolio.

issued amount, *listed*, *euro* (which are fixed characteristics of a bond), for *on-the-run* (which alters due to new issuance) and for *age* (which changes only gradually over time), we use the value of the liquidity proxy on the rebalancing date. For the proxies *missing prices*, *number of contributors* and *yield dispersion* (which depend on daily market information), we use the average value over the two weeks prior the rebalancing date. For the proxy *yield volatility* (which also depends on daily information), we calculate the standard deviation of the observed yields over the two weeks prior to the rebalancing date. If for a particular bond it is not possible to calculate the value of a liquidity proxy on the rebalancing date, that bond is ignored for that proxy until the next rebalancing date.

4. Data

The data are downloaded from three different sources. Lehman Brothers provides the International Securities Identification Numbers (ISINs) of the members of their Euro-Aggregate Corporate Bond index. The required characteristics of these corporate bonds are downloaded from Bloomberg. Reuters 3000 EXtra provides daily bid yields of each bond quoted by different pricing sources. The download period starts on 1 January 1999 and ends on 31 May 2001. The ISINs are obtained for 31 May 2000. The total number of bonds on this date equaled 1190. All bonds that are issued in euros directly after the currency's introduction are included in this analysis.

Moreover, the yield time series of each corporate bond has at least twelve months history.

4.1. Lehman Brothers

Lehman Brothers provides the ISINs of the corporate bonds in their Euro-Aggregate Corporate Bond index. This index serves as a proxy for the investment-grade euro-denominated, corporate bond market. Lehman Brothers imposes a number of criteria before the corporate bonds can enter its index. All bonds must be denominated in euros or in one of the legacy currencies. Further, all bonds are investment grade, have a fixed-rate coupon, at least one-year to maturity and an issued amount of at least 150 million euro. The country of issuance and the country of the issuer are no index criteria. The credit ratings of all corporate bonds are also provided by Lehman Brothers. All ratings are downloaded for 31 May 2000. Due to data limitations, we have kept these ratings unchanged during the whole sample period. Finally, their Euro-Aggregate Corporate Bond BBB sub index is used to construct the Fama–French credit factor.

4.2. Bloomberg

Bloomberg provides the required bond characteristics. Using the ISINs that are given by Lehman Brothers these characteristics are downloaded. In case an ISIN code is not recognized by Bloomberg, the bond data are obtained from Lehman Brothers' PC Product system. From the initial 1190 ISINs, three are not available in the Bloomberg data base. The downloaded corporate bond characteristics are: issued amount, issue date, maturity date, currency, call dates, put dates and sinking fund dates. Euro-denominated par swap data, which are used to calculate the two Fama–French factors and the portfolio excess yields, are also downloaded from Bloomberg.

4.3. Reuters

Reuters 3000 EXtra provides the bid yields of the selected corporate bonds. Most corporate bond yields in the Lehman Brothers Euro-Aggregate index are bid yields; only newly issued corporate bonds have ask yields during their first month in the index; see Lehman Brothers (1998). Therefore, we download bid yields from Reuters. For each corporate bond, all pricing sources (also called contributors) are downloaded. We exclude two Reuters pricing sources, the clearing agency ISMA and two anonymous pricing sources from the list of contributors, since they are averages of other pricing sources. The total number of different pricing sources thus obtained equals 74.

From the original 1190 ISINs in the Lehman Brothers Euro-Aggregate Corporate Bond index, 191 bonds cannot be analyzed, because they either have no Reuters Identification Code (RIC) that matches their ISIN or they do have a RIC but no contributor. For the remaining 999 bonds, all bid yields from all pricing sources are downloaded. This means that a number of time series, equal to the number of pricing sources, shows the yield development of each bond. Most bonds are quoted by more than one pricing source.

5. Results

We first present the results of applying the Fama–French bond-market model to the entire sample and show the extension of this model with portfolio characteristics. Next, the regression results for models 1 and 2 are given. Finally, the performance of the liquidity proxies is compared.

5.1. Entire sample

We first test whether the two-factor Fama–French model can be used to describe the average excess yield of all bonds in our sample. This test is relevant, because Fama and French (1993) applied their model to realized returns of US bonds, while we analyze yields of euro-dominated bonds. We estimate the following model:

$$Y_t = \alpha + \sum_{j=1}^{2} \beta_j F_{jt} + \varepsilon_t, \ \ \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2), \tag{5}$$

where the excess yield Y_t is the average bond yield (calculated over all bonds in the sample) minus the one-year euro swap rate, the slope factor F_{1t} is defined as the 10-year swap rate minus the one-year swap rate of the previous day and the credit factor F_{2t} is calculated as the Lehman Brothers Euro-Aggregate Corporate Bond BBB sub-index minus the 10-year euro swap rate.

The first row of Table 3 shows the R^2 and the estimated coefficients along with their t-values. The R^2 value is high and comparable to the values reported by Fama and French (1993). The estimated slope and credit coefficients have the expected positive sign and are highly statistically significant. The intercept is not statistically significant, so that the Fama–French model cannot be rejected for the entire sample.

To test our choice for approximating default-free interest rates with swap rates, regression model (5) is estimated again, but with swap rates replaced by government

Table 3
Results for the entire sample[a]

	Intercept	Slope	Credit	R^2 (%)
Swap rates	0.0371	0.785	0.173	97.9
	(1.01)	(36.5)	(6.66)	
Government rates	0.419	0.540	0.273	95.0
	(12.4)	(31.4)	(5.46)	

[a] Regression results for the Fama–French model estimated from the entire sample with either swap rates or government rates as default-free interest rates (t-values between parentheses).

rates. So, the excess yields and the slope and credit factors are now calculated with government yields. Our proxy for euro government rates is the Lehman Brothers Euro-Aggregate Treasury index. The second row of Table 3 shows the regression results. Both the R^2 and the t-values of the slope and credit factors have decreased compared to the model with swap rates. Moreover, the intercept is now significantly different from zero. Therefore, the Fama–French model should be rejected in case government rates are used as default-free rates. This empirically confirms our choice for using swap rates as proxy for default-free interest rates instead of Treasury rates.

5.2. Characteristics

As recommended by Gebhardt et al. (2001), we analyze the added value of incorporating characteristics into the model. We consider two characteristics:

- *Rating*: rating of the bond's issuer at 31 May 2000: AAA, AA, A or BBB.
- *Maturity*: the remaining time-to-maturity of a bond, measured in years.

To determine whether a characteristic is important for explaining excess bond yields, we follow the same procedure as for our liquidity proxies, as described in Section 3.2, except that the null model is now the Fama–French model of the previous section. For each characteristic i, we create portfolios and estimate the following regression model:

$$Y^i_{pt} = \alpha^i_p + \sum_{j=1}^{2} \beta^i_{jp} F_{jt} + \varepsilon^i_{pt},$$

where the assumptions on the disturbances are equal to those in Eq. (1). For the characteristic *rating*, we create four portfolios: portfolio 1 contains the AAA-rated bonds, portfolio 2 the AAs, portfolio 3 the As, and portfolio 4 the BBBs. For the characteristic *maturity*, two portfolios are constructed: portfolio 1 consists of the 50% shortest bonds, and portfolio 2 of the 50% longest bonds. [12]

The regression results are reported in Table 4. For *rating*, we find that the intercepts are larger for lower ratings, although the step from AA to A is very small. All Fama–French factor loadings are significant. The Wald test indicates that the four intercepts are highly jointly significant. For *maturity*, the intercepts of the portfolios reveal that short-maturity bonds have smaller yields than long-maturity bonds, with an average difference of 38 bps. The null hypothesis that the two intercepts are jointly equal to zero is easily rejected.

From these results, we conclude that the *rating* and *maturity* characteristics are important determinants of excess yield in the euro corporate bond market. To make the characteristics operational, we have to transform them to a numerical scale:

[12] The portfolios are updated every two weeks, just like in Section 3.2.

Table 4
Results for the characteristics portfolios[a]

	Intercept	Slope	Credit	Wald	R^2 (%)
Rating					
AAA	−0.220	0.736	0.0838	946	97.2
	(8.30)	(46.8)	(4.40)	(0.00)	
AA	0.120	0.732	0.0310		
	(5.60)	(55.8)	(2.04)		
A	0.122	0.856	0.295		
	(4.95)	(59.4)	(15.8)		
BBB	0.453	0.824	0.435		
	(8.28)	(24.4)	(10.6)		
Maturity					
Short	−0.135	0.635	0.165	474	98.8
	(3.09)	(25.0)	(5.36)	(0.00)	
Long	0.247	0.944	0.138		
	(13.6)	(82.0)	(10.6)		

[a] Regression results for the Fama–French model estimated from portfolios based on the rating and maturity characteristics (t-values between parentheses). The *Wald* column shows the test on the joint significance of the intercepts (p-value between parentheses).

- *Rating*: the letters are mapped as follows: AAA = 1, AA = 2, A = 3 and BBB = 4. Although this linearity assumption is somewhat crude, it is not uncommon in the literature. Moreover, since our bonds are all investment grade, and the non-linearities in S&P's and Moody's rating scales are especially apparent for speculative grade ratings, we believe that the linear scale is a reasonable approximation.
- *Maturity*: this is already a continuous variable, and thus needs no transformation.

The value of characteristic j for the pth proxy-i portfolio on day t, denoted C^i_{jpt} in Section 3.2, is calculated analogously to the liquidity variable L^i_{pt} below Eq. (2). For instance, for the characteristic *maturity*, C^i_{jpt} is the average maturity of all quoted bonds in the pth proxy-i portfolio on day t, in deviation from the average maturity of all quoted bonds on day t.

5.3. Summary statistics and correlations

For the first regression model, Eq. (1), we create two portfolios for all nine liquidity proxies. In this section, we show some summary statistics for these 18 portfolios and the correlation between the proxies.

Table 5 contains several statistics, averaged over the full sample period of 602 trading days. We observe that the average yields and average liquidity proxies for portfolio 1 (containing the hypothesized liquid bonds) and portfolio 2 (illiquid

Table 5
Portfolio statistics $P = 2$[a]

	Yield[b]		Maturity[c]		Rating[d]		Liquidity[e]	
	1	2	1	2	1	2	1	2
Issued amount	5.33	5.09	6.47	4.64	2.20	2.10	0.65	0.20
Listed	5.26	5.01	5.68	5.11	2.27	1.66	1.00	0.00
Euro	5.28	5.13	5.94	5.14	2.20	2.08	1.00	0.00
On-the-run	5.44	5.15	6.27	5.38	2.58	2.03	1.00	0.00
Age	5.44	5.16	6.91	5.31	2.42	2.09	0.64	3.80
Missing prices	5.28	5.07	6.09	4.57	2.18	2.10	0.19	0.46
Yield volatility	5.21	5.21	6.11	5.06	2.10	2.20	0.06	0.10
Number of contributors	5.19	5.27	5.58	5.56	2.13	2.19	2.31	0.76
Yield dispersion	5.40	5.14	7.42	4.91	2.18	2.13	0.47	1.50

[a] Summary statistics of the two constructed portfolios using the nine liquidity indicators. Portfolio 1 (respectively 2) contains the bonds that are hypothesized to be most liquid (respectively most illiquid).
[b] Average portfolio yield.
[c] Average time-to-maturity in years.
[d] Average credit worthiness, measured on the following scale: AAA = 1, AA = 2, A = 3, BBB = 4.
[e] Average value of the liquidity proxy.

bonds) are quite different. The yield deviations range from −29 bps (for *on-the-run*) to 8 bps (for *number of contributors*). Except for the latter proxy, we could prematurely conclude that the liquidity premium is negative, since portfolio 1 has a higher average yield than portfolio 2. However, it is not correct to fully attribute the yield differences to differences in liquidity, since the average maturity and the average rating also vary. Therefore, this table illustrates the necessity of correcting for differences in maturity and rating.

To display the correlation between the proxies, we cannot calculate the "normal" correlation (i.e. the Pearson correlation), because 5 out of 9 proxies do not change in value (or only slightly) during the sample period: *amount outstanding*, *listed* and *euro* do not change at all; *on-the-run* alters only at new issuances; *age* changes only gradually. What matters for our proxies, is that their values are used to group the bonds into a set of liquid bonds and a set of illiquid bonds. Therefore, we propose an alternative approach to display the "correlation" between two proxies. For each proxy i, we create two portfolios as described in Section 3.2. On each day t, we calculate the average value of proxy i over all bonds in portfolio 1 (and denote it by L^i_{1t}) and similarly for all bonds in portfolio 2 (resulting in L^i_{2t}). Then, for a second proxy j ($\neq i$), we calculate the average value of proxy j over all bonds in portfolio 1 (denoted by L^{ij}_{1t}) and over all bonds in portfolio 2 (L^{ij}_{2t}). If proxies i and j are both capable of splitting the sample in liquid and illiquid bonds, $L^i_{1t} - L^i_{2t}$ and $L^{ij}_{1t} - L^{ij}_{2t}$ should have the same sign. Hence, if we count the number of days that this is the case, and divide by the total number of days in our sample, we get a measure of correlation between proxies i and j. This "correlation" statistic ranges from 0% to 100%. If it equals 100%, the proxies always result in the same ordering of portfolios 1 and 2, which may be interpreted as perfect positive correlation. However, if this statistic is equal to 0%, then the ordering of portfolios 1 and 2 is always reversed, which could be described as

Table 6
"Correlation statistics" $P = 2$[a]

	Issued amount (%)	Listed (%)	Euro (%)	On-the-run (%)	Age (%)	Missing prices (%)	Yield volatility (%)	Number of contributors (%)	Yield discrepancy (%)
Issued amount		100	100	90	91	100	77	100	100
Listed	100		98	96	100	100	10	90	66
Euro	100	98		93	92	97	50	97	98
On-the-run	100	97	94		100	98	62	100	82
Age	100	92	100	100		100	91	100	97
Missing prices	100	91	95	91	96		78	100	88
Yield volatility	65	21	48	65	84	65		77	96
Number of contributors	100	75	78	80	100	100	95		79
Yield discrepancy	100	94	98	96	84	96	95	97	

[a] "Correlation" statistics between the values of the liquidity proxies for the two constructed portfolios using the nine liquidity indicators.

perfect negative correlation. Note that the statistic is not necessarily symmetric: the "correlation" between i and j is calculated on portfolios constructed with proxy i, while the "correlation" between j and i uses portfolios constructed with proxy j.

Table 6 displays the calculated statistics for all pairs. We observe that the "correlations" are remarkably high: 23 out of 72 proxy pairs are even equal to 100%, while another 29 are between 90% and 100%. In fact, only 3 "correlations" are below 50%: (*listed*, *yield volatility*) equals 10%, (*yield volatility*, *listed*) equals 21%, and (*yield volatility*, *euro*) equals 48%. Since 2 out of 3 involve the proxy *listed*, for which we do not find significant results (see Table 7 below), these low "correlations" do not worry us. The vast majority of the "correlations" are substantially above 50%, implying positive relationships between the proxies.

5.4. Model 1

Table 7 displays the results of estimating model 1 for all liquidity proxies; recall from Section 3.2 that the Fama–French factors have portfolio-specific coefficients and the characteristics common coefficients. All Fama–French factor loadings are statistically significant and have the expected positive sign. The same holds for the coefficients of the *rating* and *maturity* characteristics (with one exception: the *rating* coefficient for *issued amount* is insignificant at a 95% confidence level). All R^2-values are around 98%.

Except for the liquidity proxy *listed*, all intercept pairs are jointly statistically different from zero at a 95% significance level, as evidenced by the p-values of the Wald statistics. This indicates that the remaining eight proxies are indeed able to separate the bonds in our data set into two mutually exclusive portfolios that have statistically different yields, after controlling for differences in interest rate and credit risk. Next we look at the portfolio intercepts themselves. If our hypotheses on the sign of the

Table 7
Results for model 1[a]

	Intercept	Factors		Characteristics		Wald[b]	Premium[c]	R^2 (%)
		Slope	Credit	Rating	Maturity			
Issued amount								
Large	−0.0846	0.869	0.212					
	(2.807)	(53.1)	(10.7)	0.0850	0.127	73.5	22.9	98.0
Small	0.145	0.708	0.130	(1.76)	(8.67)	(0.00)	(0.00)	
	(4.28)	(36.3)	(5.42)					
Listed								
Yes	0.0190	0.782	0.164					
	(0.476)	(44.8)	(7.87)	0.132	0.110	2.42	−5.20	98.0
No	−0.0330	0.749	0.145	(2.16)	(7.61)	(0.30)	(0.12)	
	(0.842)	(43.2)	(6.91)					
Euro								
Euro	−0.0322	0.843	0.196					
	(1.09)	(47.6)	(8.81)	0.369	0.0727	86.0	13.6	98.0
Legacy	0.104	0.727	0.119	(7.82)	(4.94)	(0.00)	(0.00)	
	(3.55)	(41.1)	(5.66)					
On-the-run								
On	−0.0177	0.858	0.254					
	(0.648)	(52.2)	(11.8)	0.311	0.165	37.5	17.7	98.2
Off	0.159	0.771	0.145	(6.52)	(14.0)	(0.00)	(0.00)	
	(4.56)	(43.5)	(6.81)					
Age								
<1y	−0.0326	0.902	0.242					
	(1.18)	(57.0)	(11.7)	0.201	0.120	51.6	18.8	98.1
>1y	0.156	0.751	0.190	(6.60)	(12.7)	(0.00)	(0.00)	
	(5.21)	(44.9)	(8.76)					
Missing prices								
Few	−0.0329	0.798	0.146					
	(1.06)	(45.3)	(7.03)	0.195	0.163	28.4	13.1	97.6
Many	0.0984	0.732	0.179	(5.91)	(22.7)	(0.00)	(0.00)	
	(2.88)	(35.3)	(6.84)					
Yield volatility								
Small	−0.0285	0.794	0.156					
	(0.976)	(43.9)	(7.96)	0.235	0.134	11.1	13.0	97.2
Large	0.101	0.774	0.194	(4.37)	(26.6)	(0.00)	(0.00)	
	(2.58)	(32.8)	(6.91)					
Number of contributors								
Large	0.0182	0.786	0.164					
	(0.600)	(44.3)	(7.40)	0.209	0.139	53.3	13.7	97.6
Small	0.155	0.748	0.108	(7.09)	(26.2)	(0.00)	(0.00)	
	(4.98)	(40.2)	(4.82)					

Table 7 (*continued*)

	Intercept	Factors		Characteristics		Wald[b]	Premium[c]	R^2 (%)
		Slope	Credit	Rating	Maturity			
Yield dispersion								
Small	−0.0251	0.884	0.160					
	(0.82)	(48.2)	(7.56)	0.322	0.0910	53.1	22.5	98.4
Large	0.200	0.708	0.122	(8.06)	(15.9)	(0.00)	(0.00)	
	(6.71)	(39.2)	(5.69)					

[a] Regression results for the Fama–French model augmented with portfolio characteristics (see Eq. (1)) estimated from two portfolios based on one of the nine liquidity proxies (t-values between parentheses).
[b] Test on the joint significance of the intercepts (p-value between parentheses).
[c] Difference between the portfolio intercepts in basis points (p-value between parentheses).

liquidity effects are correct, the intercept of portfolio 1 should be smaller than that of portfolio 2 for all liquidity proxies. We see that this holds for eight out of nine cases; for *listed* the order is reversed, but this poses no problem, since the Wald test already indicated that for this proxy there are no significant liquidity effects.

Another way of looking at the intercepts, is to calculate their differences $\alpha_2^i - \alpha_1^i$, which we interpret as the liquidity premium for proxy i. The significance of a premium is tested with a Wald test with null hypothesis H_0: $\alpha_2^i - \alpha_1^i = 0$; the test statistic is asymptotically χ^2-distributed with 1 degree of freedom. The second to last column of Table 7 shows that the premiums for proxies *amount outstanding* and *yield dispersion* are the largest with 22.9 and 22.5 bps, respectively, while the premiums for the other proxies are between 13.0 and 18.8 bps. All premiums are statistically significant at the 95% confidence level.

5.5. Model 2

For model 2, we create four portfolios since it maximizes the power of the test for the presence of liquidity effects; see Section 3.2. Unfortunately, this means we cannot conduct the test for proxies *listed*, *age*, *euro* and *on-the-run* since they are all binary variables ('listed' versus 'not listed'; 'young' versus 'old'; 'euro' versus 'legacy'; 'on-the-run' versus 'off-the-run'). The summary statistics for the other five proxies are shown in Table 8. Clearly, the differences between the portfolios are now larger than in Table 5, since we have assigned the bonds to four size percentiles instead of two.

The regression results are displayed in Table 9. [13] The Wald statistic that tests for the joint significance of the intercept and the coefficient of the liquidity proxy is statistically significant for all five proxies. So, also using model 2, we find statistical evidence of the presence of liquidity effects in our data set. The signs of the liquidity coefficients are positive for four out of five proxies, with *number of contributors* as only (statistically insignificant) exception.

[13] The Fama–French factor loadings and the coefficients for the portfolio characteristic are omitted from Table 9 for space considerations.

Table 8
Portfolio statistics $P=4$[a]

	Yield[b]				Maturity[c]				Rating[d]				Liquidity[e]			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Issued amount	5.38	5.28	5.08	5.10	6.93	5.93	4.93	4.34	2.17	2.23	1.98	2.21	0.88	0.38	0.24	0.16
Missing prices	5.35	5.21	5.07	5.12	6.53	5.63	4.57	4.67	2.21	2.14	2.08	2.22	0.05	0.34	0.45	0.53
Yield volatility	5.25	5.18	5.17	5.26	6.71	5.65	5.25	4.84	2.11	2.08	2.10	2.31	0.06	0.07	0.08	0.12
Number of contributors	5.27	5.10	5.25	5.32	6.24	4.92	5.59	5.41	2.18	2.08	2.15	2.42	3.48	1.13	0.80	0.57
Yield dispersion	5.43	5.37	5.19	5.07	7.94	6.88	5.44	4.30	2.16	2.20	2.14	2.12	0.35	0.60	0.92	2.15

[a] Summary statistics of the four constructed portfolios using five liquidity indicators. Portfolio 1 (respectively 4) contains the bonds that are hypothesized to be most liquid (respectively most illiquid).

[b] Average portfolio yield.

[c] Average time-to-maturity in years.

[d] Average credit worthiness, measured on the following scale: AAA = 1, AA = 2, A = 3, BBB = 4.

[e] Average value of the liquidity proxy.

Table 9
Results for model 2[a]

	Intercept	Liquidity	Wald[b]	R^2 (%)
Issued amount	0.0316	0.427	259	97.9
	(1.32)	(15.9)	(0.00)	
Missing prices	0.0894	0.343	60.8	96.3
	(3.23)	(7.57)	(0.00)	
Yield volatility	0.0522	1.07	18.4	96.7
	(2.30)	(3.43)	(0.00)	
Number of contributors	0.102	−0.0037	16.8	96.2
	(3.99)	(0.618)	(0.00)	
Yield dispersion	0.0855	0.0318	27.5	97.8
	(4.26)	(3.01)	(0.00)	

[a] Regression results for the Fama–French model augmented with portfolio characteristics and a liquidity variable (see Eq. (2)) estimated from four portfolios based on one of five liquidity proxies (t-values between parentheses). The coefficients and t-values of the Fama–French factors and the characteristics are omitted for space considerations.

[b] Test on the joint significance of the intercept and the coefficient of the liquidity variable (p-value between parentheses).

Table 10
Results of the comparison tests[a]

	Adds power	Subsumes	Total
Issued amount	3	1	4
Missing prices	1	2	3
Yield volatility	2	3	5
Number of contributors	2	3	5
Yield dispersion	1	2	3

[a] Results of the pair wise comparisons of five liquidity proxies (see Eq. (3)). The table displays the number of times a proxy *adds explanatory power* to another proxy and the number of times a proxy *subsumes another proxy*.

5.6. Comparison

Table 10 summarizes the results of conducting the pairwise comparisons between the liquidity proxies, as described in Section 3.3. For each proxy i, we count the number of times it adds power to a model that already contains proxy j. We also count the number of times a proxy j is subsumed if it is added to the model of proxy i. Looking at the sum of both counts, we see that proxies *yield volatility* and *number of contributors* perform somewhat better than the other four proxies. Since the differences are small, the test does not yield a clear winner.

6. Summary

In this paper, we used the Brennan and Subrahmanyam (1996) methodology to test whether bond market liquidity is priced based on liquidity proxies: *issued amount*, *listed*, *age*, *missing prices*, *yield volatility*, *number of contributors*, *yield dispersion*, *euro* and *on-the-run*. For each liquidity proxy, we constructed mutually exclusive portfolios. The time series of portfolio yields were subsequently used in two Fama and French (1993) regression models, augmented with portfolio characteristics as recommended by Gebhardt et al. (2001), to control for differences in interest rate risk, credit risk, maturity and rating between the portfolios. We also conducted pairwise comparisons of the liquidity proxies, as proposed by Goldreich et al. (2002).

The results indicated that the null hypothesis of no liquidity premium should be rejected for eight out of nine liquidity proxies. The premium between liquid and illiquid portfolios depended on the liquidity proxy and ranged from 13 to 23 basis points. The highest premiums were found for the proxies *amount outstanding* and *yield dispersion*. The pairwise comparison tests point out that no proxy stands out from the rest.

Acknowledgements

This paper is a substantially revised version of our earlier paper 'How To Measure Corporate Bond Liquidity?'. We thank Martijn van der Voort, two anonymous referees and participants of the 2003 European Finance Association conference in Glasgow for their constructive comments. Views expressed in the paper are the authors' own and do not necessarily reflect those of Robeco Asset Management, AEGON Asset Management or ABN Amro.

References

Alexander, G.J., Edwards, A.K., Ferri, M.G., 2000. The determinants of trading volume of high-yield corporate bonds. Journal of Financial Markets 3, 177–204.

Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance 23 (4), 589–609.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid–ask spread. Journal of Financial Economics 17, 223–249.

Amihud, Y., Mendelson, H., 1991. Liquidity, maturity, and the yields on US Treasury securities. Journal of Finance 46 (4), 1411–1425.

Annaert, J., De Ceuster, M.J.K., 1999. Modelling European credit spreads. Working Paper, University of Antwerp-UFSIA. Available from <http://fetew.rug.ac.be/fineco/jema.html>.

Boudoukh, J., Whitelaw, R.F., 1991. The benchmark effect in the Japanese government bond market. Journal of Fixed Income 1 (2), 52–59.

Boudoukh, J., Whitelaw, R.F., 1993. Liquidity as a choice variable: A lesson from the Japanese government bond market. Review of Financial Studies 6 (2), 265–292.

Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373.

Brennan, M.J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41 (3), 441–464.

Chakravarty, S., Sarkar, A., 1999. Liquidity in US fixed income markets: A comparison of the bid–ask spread in corporate, government and municipal bond markets. Working Paper, Purdue University and Federal Reserve Bank of New York. Available from <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=163139>.

Chordia, T., Roll, R., Subrahmanyam, A., 2001. Market liquidity and trading activity. Journal of Finance 56 (2), 501–530.

Chordia, T., Subrahmanyam, A., Anshuman, V.R., 2001. Trading activity and expected stock returns. Journal of Financial Economics 59, 3–32.

Collin-Dufresne, P., Goldstein, R.S., Martin, J.S., 2001. The determinants of credit spread changes. Journal of Finance 56 (6), 2177–2207.

Cornell, B., 1992. Liquidity and the pricing of low-grade bonds. Financial Analysts Journal 48 (1), 63–68.

Crabbe, L.E., Turner, C.M., 1995. Does the liquidity of a debt issue increase with its size? Evidence from the corporate bond and medium-term note markets. Journal of Finance 50 (5), 1719–1734.

Díaz, A., Navarro, E., 2002. Yield spread and term to maturity: Default vs. liquidity. European Financial Management 8 (4), 449–477.

Dimson, E., Hanke, B., 2001. The expected illiquidity premium: Evidence from equity index-linked bonds. Working Paper, London Business School. Available from <http://phd.london.edu/bhanke>.

Elton, E.J., Green, T.C., 1998. Tax and liquidity effects in pricing government bonds. Journal of Finance 53 (5), 1533–1562.

Elton, E.J., Gruber, M.J., Agrawal, D., Mann, C., 2002. Factors affecting the valuation of corporate bonds. Working Paper, Stern School of Business, New York University. Available from <http://pages.stern.nyu.edu/eelton/Research.htm>.

Ericsson, J., Renault, O., 2001. Liquidity and credit risk. Working paper, McGill University and Universite Catholique de Louvain. Available from <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=248368>.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33 (1), 3–56.

Fisher, L., 1959. Determinants of the risk premiums on corporate bonds. Journal of Political Economy 67, 217–237.

Fleming, M.J., 2002. Are larger Treasury issues more liquid? Evidence from bill reopenings. Journal of Money, Credit, and Banking 3 (2), 707–735.

Fridson, M.S., Jónsson, J.G., 1995. Spread versus Treasuries and the riskiness of high-yield bonds. Journal of Fixed Income 5 (3), 79–88.

Garman, M., 1976. Market microstructure. Journal of Financial Economics 3, 257–275.

Gebhardt, W.R., Hvidkjaer, S., Swaminathan, B., 2001. The cross-section of expected corporate bond returns: Betas or characteristics? Working Paper, Axia Energy Europe, University of Maryland and Cornell University. Available from <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=281209>.

Gehr, A.K., Martell, T.F., 1992. Pricing efficiency in the secondary market for investment-grade corporate bonds. Journal of Fixed Income 2 (3), 24–38.

Goldreich, D., Hanke, B., Nath, P., 2002. The price of future liquidity: Time-varying liquidity in the US Treasury market. Working Paper, London Business School, Institute of Finance and Accounting. Available from <http://phd.london.edu/bhanke/>.

Golub, B., Tilman, L., 2000. No room for nostalgia in fixed income. Risk (July), 44–48.

Greene, W.H., 2000. Econometric Analysis, fourth ed. Prentice Hall, Englewood Cliffs, NJ.

Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401–439.

Hong, G., Warga, A., 2000. An empirical study of bond market transactions. Financial Analysts Journal 56 (2), 32–46.

Jankowitsch, R., Mösenbacher, H., Pichler, S., 2002. Measuring the liquidity impact on EMU government bond prices. Working Paper, Vienna University, CCEFM and Bank Austria. Available from <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=302795>.

Jarrow, R.A., Turnbull, S.M., 1995. Pricing derivatives with credit risk. Journal of Finance 50 (1), 53–85.

Kamara, A., 1994. Liquidity, taxes, and short-term treasury yields. Journal of Financial and Quantitative Analysis 29 (3), 403–417.

Kempf, A., Uhrig-Homburg, M., 2000. Liquidity and its impact on bond prices. Schmalenbach Business Review 52, 26–44.

Kocić, A., Quintos, C., Yared, F., 2000. Identifying the benchmark security in a multifactor spread environment. Fixed Income Derivatives Research, Lehman Brothers, New York.

Krishnamurthy, A., 2002. The bond/old-bond spread. Journal of Financial Economics 66 (2–3), 463–506.

Lehman Brothers, 1998. Introduction to the Pan-European Aggregate Index, Research document, Fixed Income Research.

Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3 (3), 413–467.

Lys, T., Sabino, J.S., 1992. Research design issues in grouping-based tests. Journal of Financial Economics 32 (3), 355–387.

McGinty, L., 2001. Issue size versus liquidity in credit. J.P. Morgan Fixed Income Research.

Mullineaux, D.J., Roten, I.C., 2002. Liquidity, labels, and medium-term notes. Financial Markets, Institutions and Instruments 11, 401–423.

Newey, W., West, K., 1987. A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Sarig, O., Warga, A.D., 1989. Bond price data and bond market liquidity. Journal of Financial and Quantitative Analysis 24 (3), 367–378.

Schultz, P., 2001. Corporate bond trading costs and practices: A peek behind the curtain. Journal of Finance 56 (2), 677–698.

Shulman, J., Bayless, M., Price, K., 1993. Marketability and default influences on the yield premia of speculative-grade debt. Financial Management 22 (3), 132–141.

Smidt, S., 1971. Which road to an efficient stock market: Free competition or regulated monopoly. Financial Analysts Journal 27 (5), 18–20 64–69.

Strebulaev, I.A., 2001. Many faces of liquidity and asset pricing: Evidence from the US treasury securities market. Working Paper, London Business School. Available from <http://facultyresearch.london.edu/detwp.asp?id=4952>.

Tychon, P., Vannetelbosch, V., 2002. Debt valuation and marketability risk. Working paper, Université Catholique de Louvain. Available from <http://www2.econ.ucl.ac.be/Users/v.vannetelbosch/>.

Vayanos, D., 1998. Transaction costs and asset prices: A dynamic equilibrium model. Review of Financial Studies 11 (1), 1–58.

Warga, A.D., 1992. Bond returns, liquidity and missing data. Journal of Financial and Quantitative Analysis 27 (4), 605–617.