Expected Bond Liquidity*

Michael Reichenbacher† Philipp Schuster‡ Marliese Uhrig-Homburg§

December 21, 2020

Abstract

We introduce an approach to forecast individual bond liquidity and apply it to the U.S. corporate bond market. Our model combines three dynamic prediction models to get the most accurate estimate for future bond liquidity. We compare the new prediction methodology with the literature's current approach to use a bond's liquidity of today as the best estimate for its liquidity tomorrow. Our approach generates significantly lower forecasting errors and is much better able to capture the premium for expected liquidity in bond yields. We provide evidence that investors in corporate bond funds actively anticipate liquidity deterioration in underperforming funds and sell their shares in advance to secure a first-mover advantage.

JEL classification: C10, C53, G12, G17, G20
Keywords: bond liquidity, bid-ask spread, forecasting, asset pricing, bond funds

*We thank Nina Boyarchenko, Martin Ruckes, and Julian Thimme for helpful comments and suggestions. In addition, we thank the DZ BANK Foundation for sponsoring our Bloomberg professional account and Deutsche Forschungsgemeinschaft for data access through CRC 649 Economic Risk. This work was supported by the Deutsche Forschungsgemeinschaft [UH 107/3-2 and SCHU 3049/2-2].

† Karlsruhe Institute of Technology, Institute for Finance, P.O. Box 6980, D-76049 Karlsruhe. Email: michael.reichenbacher@kit.edu.

‡ University of Stuttgart, Department of Finance, Keplerstr. 17, D-70174 Stuttgart. Email: philipp.schuster@bwi.uni-stuttgart.de.

§ Karlsruhe Institute of Technology, Institute for Finance, P.O. Box 6980, D-76049 Karlsruhe. Email: uhrig@kit.edu.

1 Introduction

A basic principle of financial economics is that expectations about future market conditions influence decisions. Because liquidity is an elusive concept and the literature has found multifaceted relations with other market factors, forming expectations about future liquidity is difficult. In the absence of a generally accepted forecasting approach, market participants are left alone to aggregate and extrapolate the available information when they assess future liquidity. For example, investors require the liquidity at the future time of sale to evaluate the expected payoff of a trading strategy. Regulators and central banks monitor the expected development of liquidity very closely to take timely countermeasures. Finally, issuing companies react to their bonds' expected liquidity deteriorations to avoid distress arising from worsening refinancing conditions (see, e.g., He and Xiong, 2012).

We are not aware of a (sophisticated) forecasting model for individual bond liquidity in the academic literature.[1] Instead, researchers employ the naïve assumption that a bond's liquidity today is the best estimator for its liquidity tomorrow. We fill this gap and introduce a forecasting model for individual bond liquidity. Our objective is to employ the information available to a contemporary forecaster to most precisely estimate a bond's liquidity in the month ahead. To this end, we exploit the large pool of drivers of liquidity from the literature and dynamically select each month the subset of predictors that offers the best predictive power given the current information set.

Our forecasting procedure combines elements from machine learning with the transparency of a simple linear model. In each month, our algorithm performs the following steps based on a rolling window of the previous twelve months. First, we use three different approaches to select the predictor variables that have the strongest forecasting power for the most recent past. We use elastic net, a variant of stepwise regression, and a method that relies on significant relations within the calibration window. We then calibrate a simple linear model to the selected variables. To further increase the predictive accuracy, we combine the forecasts from the three selection approaches to an average forecast (see, e.g., Rapach, Strauss, and Zhou, 2010). We implement the prediction model on the U.S. corporate bond market for the simple average bid-ask spread measure of Hong and Warga (2000) using transaction data from Enhanced TRACE for the period from October 1, 2004 to June 30, 2017. Note that the procedure can be easily applied to any liquidity measure and we consider a more advanced liquidity measure in the robustness section.

[1] Forecasting liquidity at the market level, Boyarchenko, Giannone, and Shachar (2019) find that autoregressive models are hard to beat.

We evaluate the performance of the new prediction model relative to the literature's naïve approach on the basis of a direct and an indirect comparison. First, we compare the forecasting errors in an out-of-sample setting for our forecasting model and the naïve prediction. We find that our new model outperforms the naïve prediction model in every month of our observation period from 2004 to 2017. Interestingly, the largest performance improvements occur during the financial crisis. Overall, our forward-looking approach reduces the average forecasting error by about 19%. Second, in the indirect comparison, we show that the predictions of our new model better explain the premium for expected liquidity in corporate bond yields. We exploit Amihud and Mendelson's (1986) finding that investors require higher expected returns for assets that trade at higher (future) transaction costs. Following their guidance, we regress, in a panel setting, monthly yield-spread changes on changes in expected liquidity and a set of control variables. We compare the results of this analysis using expected liquidity from our forecasting model with the results when using the naïve assumption that a bond's liquidity today is the best estimator for its liquidity tomorrow. We find a much higher sensitivity of yield spreads to expected liquidity combined with a higher explanatory power when we use our sophisticated model. The about seven times higher sensitivity of yield spreads to changes in expected liquidity indicates that a naïve approach strongly underestimates the influence of liquidity on financing costs.

Finally, we leverage the forward-looking nature of our approach and shed light on the strategic behavior of investors in corporate bond funds. When selling their shares, investors in mutual funds usually receive the net asset value as of the time of sale. Costly portfolio readjustments, however, happen at a later date and lead to negative externalities for the investors who stay in the fund. The resulting first-mover advantage can lead to 'runs' on the fund similar to bank runs. Consequently, flows out of poorly performing funds are exacerbated when the funds' portfolio is illiquid (Goldstein, Jiang, and Ng, 2017). Given this background, we analyze whether investors anticipate liquidity deteriorations and incorporate this information into their redemption decisions. To discriminate between investors who react to observed liquidity and those that actively form expectations, we regress, in a panel setting, monthly corporate bond fund flows on a fund's current liquidity and its expected liquidity change. For funds with a negative performance, we find evidence consistent with investors indeed acting on expected liquidity deteriorations. This anticipation channel reinforces the established effect that investors oversell poorly performing funds with currently illiquid holdings. Intuitively, both effects become more pronounced if a fund's performance gets worse.

We do not claim that market participants have formed their expectations in exact accordance with our prediction model, and we concede that their prediction approaches may even work better. However, market participants' forecasts are likely correlated with our predictions. Indeed, our results indicate that the forecast of the marginal investor is correlated stronger with our prediction than with a bond's current liquidity as the naïve prediction. This finding contrasts the practice in the literature to employ a bond's current liquidity when, formally, expected liquidity is required. For example, in asset pricing applications on bond liquidity, essentially all papers use a bond's current liquidity (see, e.g., Bao, Pan, and Wang, 2011; Friewald, Jankowitsch, and Subrahmanyam, 2012; Dick-Nielsen, Feldhütter, and Lando, 2012; Bongaerts, de Jong, and Driessen, 2017).[2] Moreover, our robustness test using a different forecasting methodology based on a random forest model reveals that our findings do not depend on the exact procedure used to calculate forecasts.

2 Predicting Bond Liquidity

In this section we introduce our liquidity prediction approach. We identify a set of *candidate* predictor variables for which a close connection with bond liquidity has been documented in the literature. For each point in time, our forecasting algorithm selects from this set of candidates the variables that have the highest predictive power in the most recent past. Our goal is to find the best-performing model from the perspective of a contemporary observer, mitigating the impact of a look-ahead bias that would arise if the variables were selected based on full-sample information. We compare the performance of our prediction model with a naïve benchmark model assuming a bond's liquidity in the next month is unchanged from today. Such a naïve forecast is exactly what researchers implicitly do when they use the currently prevailing liquidity in their applications instead of the expected liquidity actually required. Hence, we also examine whether our measure for predicted liquidity is able to better explain changes in bond yield spreads compared to the naïve approach.

2.1 Data and Liquidity Measure

Our analysis is based on bond transaction data from Enhanced TRACE from October 1, 2004 to June 30, 2017 (see Appendix A for details). Bond characteristics, rating histories, and outstanding amounts are from Reuters Eikon and Bloomberg. We implement our liquidity

[2]Some papers argue that (market) liquidity follows an AR-1 process (see, e.g., Amihud, 2002). If an AR-1 process is the best model to describe liquidity movements, today's liquidity contains all information and could be used as proxy for tomorrow's expected liquidity. Our results indicate that this is not the case.

forecast for the commonly used average bid-ask spread measure of Hong and Warga (2000) and for a more advanced liquidity measure that incorporates the dependence of transaction costs on trade size in the robustness section. The average bid-ask spread for bond i in month t can be calculated as

$$AvgBidAsk_{i,t} = Avg\left[\frac{\overline{P_{i,d}^{buy}} - \overline{P_{i,d}^{sell}}}{0.5 \cdot \left(\overline{P_{i,d}^{buy}} + \overline{P_{i,d}^{sell}}\right)}\right], \tag{1}$$

where $\overline{P_{i,d}^{buy/sell}}$ is the average of all buy or sell trades in bond i on day d.

2.2 Drivers of Liquidity

Despite bond liquidity being persistent (see, e.g., Chordia, Sarkar, and Subrahmanyam, 2005; Acharya, Amihud, and Bharath, 2013), a large body of empirical literature shows that it varies predictably over the lifetime of a bond. Early studies find that bonds are typically most liquid directly after issuance and get more illiquid when they age (see, e.g., Warga, 1992; Hong and Warga, 2000). Bonds with a higher outstanding amount and bonds that trade more frequently have lower transaction costs (see, e.g., Edwards, Harris, and Piwowar, 2007; Bao, Pan, and Wang, 2011; Jankowitsch, Nashikkar, and Subrahmanyam, 2011). Riskier bonds with a higher duration and more credit risk are typically less liquid than comparable bonds with lower risks (see, e.g., Mahanti, Nashikkar, Subrahmanyam, Chacko, and Mallik, 2008; Hotchkiss and Jostova, 2017). Chordia, Roll, and Subrahmanyam (2000) show for the stock market that individual trading costs move together with market and sector specific trading costs. Liquidity is also related to broader measures of market functioning. For example, Chordia, Sarkar, and Subrahmanyam (2005) and Goyenko and Ukhov (2009) identify bond market performance, volatility, order imbalance, and spillover effects from the stock market as driving factors of bond market liquidity. Additionally, these authors find that macroeconomic variables such as monetary policy, inflation, or industrial production have a significant connection to liquidity.[3]

Based on this literature, we build our set of candidate predictors to forecast next period's bid-ask spreads $\widehat{AvgBidAsk}_{i,t+1}$. Given the high persistence of individual and market liq-

[3]We do not consider bond characteristics that usually do not change during a bond's life such as the coupon, the original time-to-maturity, embedded options, or industry effects (see, e.g., Edwards, Harris, and Piwowar, 2007; Hotchkiss and Jostova, 2017; Jankowitsch, Nashikkar, and Subrahmanyam, 2011; Mahanti, Nashikkar, Subrahmanyam, Chacko, and Mallik, 2008). We capture the effects of such time-invariant variables through a bond's lagged liquidity.

uidity, we naturally include a bond's bid-ask spread in the current month t. Because many bonds trade very infrequently leading to potentially noisy liquidity measures, we additionally include the moving average of the liquidity measure from the previous twelve months. As liquidity and its standard deviation are closely related (Dick-Nielsen, Feldhütter, and Lando, 2012), we also consider the standard deviation of the daily bid-ask spread measure. Next, we include a bond's age, its duration, and its (log-transformed) outstanding amount.[4] We capture trading activity with the logarithms of the average trade size and the total trading volume. Following Chordia, Sarkar, and Subrahmanyam (2005), we incorporate a bond's monthly return and order imbalance as possible predictors. We measure bond order imbalance as the difference between a bond's buying and selling dollar volume normalized with total trading volume. Regarding credit risk, we use the average numerical bond rating of the three rating agencies S&P, Moody's, and Fitch[5] and the five-year CDS spread from Markit.

Given the strong commonality of individual liquidity with market-wide liquidity, we include aggregate corporate bond market liquidity. We measure monthly market liquidity as the equally-weighted average bid-ask spread across all bonds in the sample. Before aggregating, we winsorize spreads at the 1% and 99% levels. In the spirit of Chordia, Roll, and Subrahmanyam (2000), we include a more granular aggregate liquidity measure, which is motivated by the fact that bonds with similar characteristics might be driven by the same market forces as they are to some extent substitutes to each others. To this end, we perform an independent triple sort in each month. Each portfolio represents a bond segment and we use the average portfolio bid-ask spread for each bond in the portfolio as candidate predictor. We follow Bongaerts, de Jong, and Driessen (2017) and Downing, Underwood, and Xing (2005) by sorting on a bond's average rating (quartiles), on its amount outstanding (terciles), and on its time to maturity (terciles), leading to 36 different segments.[6]

Following Goyenko and Ukhov (2009), we include short- and long-term market returns using the one-month and twelve-month return of the Barclay's U.S. corporate bond index. We also employ market volatility and order imbalance (see Chordia, Sarkar, and Subrahmanyam, 2005). We capture bond market volatility via CBOE's 10-year U.S. treasury note volatility index (TYVIX) and via the annualized realized volatility of Barclay's U.S. corporate bond index within month t. Market order imbalance is calculated as the difference

[4]Note that our data for amount outstanding from Reuters Eikon includes reopenings, repurchases, and other (early) redemptions.

[5]We transform the ratings to integer numbers (AAA: 1, ... D: 22).

[6]As a robustness check, we also test a dependent version of the triple sort in which we first sort on rating, then on amount outstanding, and last on time to maturity. Additionally, we test an alternative ordering following Downing, Underwood, and Xing (2005) and sort on time to maturity, rating, and finally on the amount outstanding. In both settings, results remain qualitatively the same.

between monthly aggregate buying and selling dollar volume normalized with total trading volume. Considering spillover effects from the equity market, we use the one-month and twelve-month return of the S&P 500. We also include equity market volatility via CBOE's volatility index (VIX) as well as equity market liquidity.[7] Following common practice, we use the Amihud (2002) price impact measure to approximate equity market liquidity.[8] Regarding broader (macro) economic factors, we include the inflation rate using data from the OECD as well as the one-month and six-month TED spread, the federal funds rate to approximate monetary policy, and yearly industrial production growth using data from the Federal Reserve Bank of St. Louis.[9]

We report descriptive statistics for our set of candidate predictor variables in Table 1. Panel A shows average cross-sectional statistics for the variables with both time-series and cross-sectional variation. The average bond has a bid-ask spread of 1.3%, a monthly return of 0.5%, an age of roughly four years, and a duration of roughly six years. Regarding trading activity, the average bond has a log-transformed total trading volume of 17.6 (corresponding to about $45 million) and exhibits a slightly negative order imbalance. Regarding credit quality, the average bond has a rating of about 8 (corresponding to BBB+) and a credit default swap spread of 1.7% p.a. Overall, all variables show a strong variation in the cross-section. Panel B reports time series statistics for our market and macroeconomic variables. For both bond and equity market liquidity, the 95%-percentile shows that during crises, market-wide trading costs and price impact have been about two times as large as on average. Corporate bond and equity markets generate an average yearly return of about 5% and 9% during our sample period. In aggregate, the bond market exhibits a slightly negative order imbalance. Regarding macroeconomics, the average inflation rate is about 2% and industrial production growths by roughly 0.7% per year. Again, all variables show a strong variation.

Insert Table 1 about here.

Table 2 presents correlations for our set of candidate predictors. We additionally include next month's liquidity to get a first insight on which variables might have predictive power. As expected, we observe a strong correlation of 0.71 between current liquidity and next

[7]We do not include stock market order imbalance due to data limitations.

[8]We calculate a monthly measure as the equally-weighted mean from all stocks with share codes of 10 and 11 in CRSP. We exclude observations on days without trading and require at least three days with positive trading volume per month in a stock. Further, we exclude shares traded at NASDAQ and winsorize the monthly cross-section of individual Amihud (2002) measures at the 5% and 95% level.

[9]Note that data on inflation and industrial production becomes available with a time lag of one month. Thus, we lag these two variables by one month.

month's liquidity. The high persistence is also confirmed by the correlation of 0.73 between next month's liquidity and the moving average of the previous twelve months. This finding shows that a naïve forecast of future liquidity using current liquidity is not a bad starting point. However, there are several other variables that also show a strong correlation with next month's liquidity. Especially the trading activity variables exhibit a promising relation. Consistent with the literature, we find that a higher average trade size and a higher total trading volume are associated with lower bid-ask spreads. Also as expected, we find a negative correlation of -0.43 between outstanding amount and next month's liquidity and confirm the well-known negative relation between credit quality and future liquidity (see He and Xiong, 2012). Consistent with our ex-ante expectation that older bonds are associated with lower liquidity, we find a positive correlation of 0.17 between age and next month's liquidity. In the same spirit, a positive correlation of 0.32 indicates that a higher duration leads to higher bid-ask spreads. Regarding market and macroeconomic variables, we find aggregate bond market and stock market liquidity to have the highest correlation with next month's liquidity of 0.36 and 0.33, respectively, confirming the co-movement between individual and market liquidity. Not surprisingly, our more granular aggregate liquidity measure (segment liquidity) with a correlation of 0.63 is even more strongly connected to a bond's liquidity in the next month. Interestingly, the sign and the magnitude of all correlations remain comparable when we consider current liquidity instead of next month's liquidity.

Insert Table 2 about here.

2.3 Prediction Model

Starting with the set of candidate predictors, we develop our forecast methodology. We implement an estimation procedure that exploits information up to time t to forecast liquidity in month $t+1$. In each month, we include only those variables in the model that increase the predictive power for the most recent past so that a contemporary forecaster would have been able to use the same information. The forecast is then based on the linear model

$$\widehat{AvgBidAsk}_{i,t+1} = \widehat{\alpha}_t + \sum_{m \in M_t} \widehat{\beta}_{m,t} \cdot \text{predictor}_{m,i,t}, \tag{2}$$

where, for each month t, we determine the set of predictor variables M_t and parameters $\widehat{\alpha}_t$ and $\widehat{\beta}_{m,t}$ based on a rolling window of twelve months $t-12, ..., t$ and a two-step procedure. In the first step, we select those variables M_t that have the highest predictive power within the twelve-month window. Therefore, M_t adapts to new information as it becomes available (see, e.g., Chinco, Clark-Joseph, and Ye, 2019; Pesaran and Timmermann, 1995, for a similar argument) and naturally accommodates structural changes in the relation between variables.[10] In the second step, we calibrate $\widehat{\alpha}_t$ and $\widehat{\beta}_{m,t}$ for the selected predictors. To mitigate the impact of outliers on variable selection and calibration, we winsorize the data at the 1% and 99% levels on a monthly basis.

We employ three different methods to determine the predictor set M_t in the first step. Two of the three selection methods are based on out-of-sample cross-validation. For these, we employ a holdout procedure in which the data of the first eleven months is used to train the model and the last month is used to validate the model's prediction performance. The predictor variables that lead to the lowest error on the validation set are then selected for M_t. Our first selection algorithm is a variant of stepwise regression (see, e.g., Agarwal and Naik, 2004; Titman and Tiu, 2011). Here, the algorithm adds variables that decrease the mean square error (MSE) on the validation set until no other variable leads to further improvement. After each addition, the algorithm checks if dropping one of the selected variables decreases the model's error. The second selection algorithm is the elastic net procedure, which uses a combined penalty function of LASSO and ridge methods (see, e.g., Kozak, Nagel, and Santosh, 2020; Panopoulou and Vrontos, 2015). Again, the parameters controlling the number of variables included in the model are set to minimize the MSE on the validation data set. The third selection method relies on detecting stable predictive relations in-sample (see, e.g., Chernobai, Jorion, and Yu, 2011). The model is estimated using all candidate predictors and M_t simply contains those variables for which we observe a predictive relation with a significance level of lower than 5%, where we cluster standard errors by bond.[11] In the second step, we then use the full twelve months of data to calibrate the model on the predictors selected by the three selection methods.

For the three different methods, we can now calculate bond i's expected liquidity in the next month $\widehat{AvgBidAsk}_{i,t+1}$ using Equation (2) and the predictors' values in month t.[12] Rapach, Strauss, and Zhou (2010) find that combining predictions of individual models often

[10] There are several alternatives to our approach when implementing a dynamic estimation procedure. First, one can use a recursive scheme instead of a rolling window. Second, the literature uses various lengths for rolling windows from twelve months up to five years (see, e.g, Fama and MacBeth, 1973; Kacperczyk, Nieuwerburgh, and Veldkamp, 2014). We find in unreported results that the higher flexibility of a rolling scheme and a window length of twelve months is more capable to adjust for structural changes and outperforms alternative specifications.

[11] Table 2 shows that many possible predictors are strongly connected. To mitigate multicollinearity issues, we exclude all variables with a variance inflation factor of 10 or higher (see, e.g., Liu and Ritter, 2011).

[12] In the rare case that $\widehat{AvgBidAsk}_{i,t+1}$ is negative, we set it to 0.

results in superior performance. We follow these authors and average the three forecasts to arrive at our final prediction of next month's liquidity.[13]

In the robustness section 4.1, we also test an alternative prediction approach based on a random forest model. Albeit the machine learning model brings a slight improvement in terms of forecast accuracy, it should be noted that the relation of predictors to next month's liquidity and their economic significance is notoriously difficult to interpret within such approaches.

2.4 Prediction Results

We compare the accuracy of our new forecast methodology to the naïve benchmark model $\widehat{AvgBidAsk}_{i,t+1}^{\text{naïve}} = AvgBidAsk_{i,t}$. As discussed before, researchers that apply today's liquidity for applications that formally require future expected liquidity implicitly employ such a naïve model. The forecast evaluation criterion is the root mean mean square error (RMMSE)

$$RMMSE^{\text{naïve/fc}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\frac{1}{n_t}\sum_{i=1}^{n_t}\left(AvgBidAsk_{i,t+1} - \widehat{AvgBidAsk}_{i,t+1}^{\text{naïve/fc}}\right)^2},$$

where $AvgBidAsk_{i,t+1}$ is the realized liquidity of bond i in month $t+1$ and $\widehat{AvgBidAsk}_{i,t+1}^{\text{naïve/fc}}$ is the prediction of the naïve benchmark or the forecast model for month $t+1$ based on data from t. n_t is the number of bonds in month t for which we can assess liquidity in month $t+1$. Finally, T is the number of months in our observation period. As we employ a rolling window of twelve months to calibrate our model, predictions start in November 2005 and end in June 2017.

Insert Table 3 about here.

The average forecasting errors of the naïve benchmark and our forecast model are reported in Panel A of Table 3. Using a bond's current liquidity as the naïve forecast leads on average to an error of 91 basis points. In comparison, our linear combination model only has a forecasting error of 74 basis points, which corresponds to a relative outperformance of roughly 19%. To get a better understanding of the superior performance of our forecasting model, we

[13]The econometric literature finds that the simple average of different forecasts often outperforms more sophisticated weighting schemes (the phenomenon is called 'Forecast Combination Puzzle', see, e.g, Smith and Wallis, 2009). Combining the information of the three models generates the best predictions for our sample. However, the individual models also generate good results.

plot the time series of the monthly root mean square error (RMSE) in Panel A of Figure 1. The performance of the new model surpasses the benchmark in each month of our observation period. Interestingly, while the highest prediction errors occur during the financial crisis, the largest improvement of the predictive accuracy also seems to coincide with the crisis period. Because this period is also the period with the highest average bid-ask spread, we plot the improvement in the RMSE of our forecasting model relative to the average bid-ask spread in Panel B of Figure 1. This figure shows that the outperformance compared to the naïve benchmark is relatively stable over time. Remarkably, the two months with the highest relative outperformance are within the financial crisis (March and December 2008). Thus, the application of our new model is especially advisable during times of liquidity stress.

Insert Figure 1 about here.

Finally, to investigate the source of our model's forecasting power, we report the predictors that our forecasting algorithm selects and their economic significance in Panel B of Table 3. The variables are ranked by their selection frequency, i.e., the percentage of months for which they are included in the model. Because our prediction model is based on three individual selection approaches, we calculate the average frequency. In the same spirit, we report the economic significance as the average across the monthly products of the predictor's average coefficient and its standard deviation in the 12-month rolling window used for calibration. As can be expected, the current liquidity and its 12-month moving average have the strongest impact on next month's forecast. Both variables are almost always included in the predictor set and a liquidity deterioration of one standard deviation is associated with an increase of next month's bid-ask spread of 30.4 and 41.8 basis points, respectively. The variables on rank 3 and 4 are duration and average trade size with an inclusion rate around 90 to 93%. A higher duration is associated with lower liquidity, while larger average trade sizes lead to more narrow spreads. The economic significance of both predictors with 6.7 and -5.3 basis points, is, however, much lower compared to the two autoregressive variables. Consistent with He and Milbradt (2014), we also find credit risk to be a strong driver of future liquidity. A bond's credit spread is in more than 80% of our observation months part of the predictor set and a one standard deviation deterioration in credit quality leads to a 3.9 basis points higher bid-ask spread. On rank 6, the 77% inclusion rate of a bond's segment liquidity emphasises that a bond's future liquidity is driven by co-movements in the bond market. A one standard deviation decrease in liquidity of a segment is associated with a bid-ask spread increase of 5.1 basis points. Comparing the importance of segment liquidity with that of market-wide liquidity, which is on the second to last rank, indicates that there are systematic differences in the liquidity dynamics between bonds of different characteristics. Finally, total trading volume (rank 7), bid-ask spread volatility (rank 8), amount outstanding (rank 9), and a bond's return (rank 10) complete the set of variables that are selected in more than 50% of the cases.

In the following empirical sections, we evaluate the performance of our liquidity forecast in an asset pricing environment and we examine the effect of expected liquidity on corporate bond fund flows. In such analyses, one faces the challenge that the predictor variables can have an indirect effect on the outcome variable that is not related to liquidity. For example, our candidate predictor set contains information on credit quality, market factors, and macroeconomic variables. To suppress any indirect effects of these variables, we employ a restricted version of our prediction model. To this end, we exclude all variables that are not directly related to liquidity. The restricted set of candidate predictors then only includes the current liquidity, the 12-month moving average of liquidity, volatility of liquidity, average trade size, total trading volume, outstanding amount, segment liquidity, and bond market liquidity.

Since an exclusion of variables with potential predictive power can lead to a loss in forecast accuracy, we check whether the superior performance of the forward-looking approach also holds for the restricted model. Panel A of Table 3 shows that the restricted model also outperforms the naïve benchmark model. Because the restricted model requires only a fraction of the input variables, the number of observations is much higher compared to the full model. With 80 basis points, our forecast model generates an error that is 19 basis points lower than the error of the naïve benchmark model. Again, this corresponds to a relative improvement of about 19%.[14]

2.5 Premia for Expected Liquidity in Bond Yields

In addition to the direct evaluation from the previous section, we take an indirect approach to evaluate our forecasting procedure based on an asset pricing analysis. For that, we exploit the prediction of Amihud and Mendelson's (1986) model that illiquid assets command higher expected returns. Because (expected) cash flows to investors depend on future transaction costs, only expected and not current liquidity matters for security prices and expected re-

[14]The very similar improvements of the restricted model and the full model over the naïve benchmark are partly due to the additional observations available when using a smaller number of predictors. When we compare the forecasting performance of both models on the same data set, the restricted model is dominated by the full model with a 0.4 basis points lower RMMSE.

turns.[15] For that reason, the literature's approach to use today's realized liquidity in asset pricing analyses is identical to employing the naïve proxy for future expected liquidity. In this spirit, Friewald, Jankowitsch, and Subrahmanyam (2012) examine the effect of changes in liquidity on the yield spread of bonds over the Treasury curve. Based on the higher predictive accuracy of our new model compared to the naïve approach, we expect to better capture changes in investors' expectation and, as a result, to better explain the liquidity premium embedded in bond prices.

We test our hypothesis within the setting of Friewald, Jankowitsch, and Subrahmanyam (2012) and perform a monthly panel regression of first differences of yield spreads on changes in expected future liquidity:

$$\begin{aligned}\Delta(\text{Yield spread})_{i,t} &= \alpha + \beta \cdot \Delta(\text{Yield spread})_{i,t-1} + \gamma \cdot \Delta(\widehat{AvgBidAsk})_{i,t+1} \\ &\quad + \delta \cdot \Delta(\text{Controls})_{i,t} + \epsilon_{i,t}, \end{aligned} \tag{3}$$

where $\Delta(\widehat{AvgBidAsk})_{i,t+1}$ is the change in predicted liquidity based either on our restricted model of Section 2.3 or the naïve approach. Note that $\Delta(\widehat{AvgBidAsk})_{i,t+1}$ contains only information available in month t.[16] Following Friewald, Jankowitsch, and Subrahmanyam (2012), we control for autocorrelation in yield spreads, credit risk, and other liquidity dimensions. The yield spread of a bond is its spread over the Treasury curve calculated via a theoretical bond with the same cash flow structure. We calculate daily yield spreads as volume-weighted average across all trades in TRACE. Bond i's yield spread in month t is then the average across all daily spreads (for more details, see Appendix A). We winsorize each month yield spread changes and changes in expected liquidity at the 1% and 99% levels. The credit risk of a bond is represented via changes in 21 rating dummies based on the average rating of the three rating agencies S&P, Moody's, and Fitch. We set the k-th rating dummy to 1 if the average rating is in the interval $[k-0.5, k+0.5)$, otherwise we set its value to 0. Finally, to control for other dimensions of liquidity, we employ monthly changes in the logarithm of the outstanding amount, changes in the number of trades, and changes in the logarithm of the average trade size.

The results of the panel regression (3) are reported in Table 4. Specification (1) only includes the autoregressive term and the control variables and serves as our baseline to

[15] Note that in Amihud and Mendelson's (1986) model, transaction costs are constant so that there is no difference between today's realization and future expectations of liquidity. The model can be easily extended to allow for varying transaction costs over time.

[16] For the naïve approach, $\Delta(\widehat{AvgBidAsk})_{i,t+1} = AvgBidAsk_{i,t} - AvgBidAsk_{i,t-1}$ (as $\widehat{AvgBidAsk}^{\text{naïve}}_{i,t+1} = AvgBidAsk_{i,t}$).

evaluate the impact of an inclusion of expected liquidity. We find that yield spreads show a positive autocorrelation that is significant at the 10% level.[17] Consistent with intuition, we find that a higher average trade size is associated with a significantly lower yield spread. For the number of trades, the coefficient is counterintuitively positive. Friewald, Jankowitsch, and Subrahmanyam (2012) attribute this result to investors splitting their trades in an illiquid market. In specification (2), we add a bond's current liquidity as the naïve proxy for expected liquidity to the regression model. Consistent with previous findings, we see a highly significant positive effect. An increase of the bid-ask spread by 1% is associated with a yield spread increase of about 6 basis points. However, the additional explanatory power is rather moderate with an absolute increase in the R^2 of 0.0033 compared to specification (1).[18] If we instead employ our forecasting model in specification (3), we find that the size of the coefficient increases by a factor of about seven. An increase in the expected bid-ask spread of 1% is now associated with an increase in the yield spread of about 45 basis points. The effect on the explanatory power is also more pronounced. The R^2 increases by 0.0184 compared to specification (1), which is more than five times the increase when employing the naïve benchmark.

Insert Table 4 about here.

We test whether this increase in explanatory power can be verified in an out-of-sample setting. To this end, we estimate implied yield spread changes where we employ a backward-looking rolling window of 24 months (with at least 12 months at the start of the observation period) to calibrate the regression model (3). We then calculate the mean square error between implied and actual changes. To test whether the resulting MSE of specifications (1) to (3) are significantly different, we employ a Diebold and Mariano (1995) test in the spirit of Harvey, Leybourne, and Newbold (1997).[19] Consistent with the in-sample findings, the MSE of 0.942 of the baseline model decreases significantly by 0.003 when adding the current liquidity as the naïve benchmark to the regression. For the specification in which the forecasts are based on our linear combination model, this decrease is much stronger with

[17]The positive autocorrelation is consistent with Duffee (1998). In contrast, Friewald, Jankowitsch, and Subrahmanyam (2012) find that the autocorrelation is significantly negative based on a weekly analysis. When we switch from a monthly to a weekly frequency, we also get a negative estimate, which is probably due to microstructure noise that is more important for higher frequencies.

[18]The increase in the R^2 is comparable to the results of Friewald, Jankowitsch, and Subrahmanyam (2012). These authors observe an increase of 0.009 when adding four different liquidity measures simultaneously.

[19]Note that, because of the rolling estimation, the test corresponds to an unconditional Giacomini and White (2006) test, which takes uncertainty in the parameter estimation into account. This test is equally suited to compare the nested and nonnested models of specifications (1) to (3).

0.024. Lastly, we test whether the mean square errors of the two models are significantly different. And indeed the decrease of 0.021 compared to the benchmark proxy is highly significant. Thus, the new model offers an out-of-sample improvement that is more than eight times the improvement of the naïve model.

Summarizing, our prediction model outperforms the literature's approach to measure expected liquidity using the current liquidity of a bond. We verify the superior performance using two independent analyses. First, we directly compare the predictions with future realizations. Second, our model's forecasts are able to better explain yield spread changes. Notably, the effect of a given bid-ask spread change on bond yield spreads is about seven times larger using the forward-looking approach compared to what is standard in the literature. It is interesting that the 19% lower forecasting error leads to such a strong difference in the relation of expected liquidity with yield spread changes. There are two possible channels that can explain why our forecasts perform so much better compared to realized liquidity. First, much of the variation in liquidity is probably not predictable as it depends on new information becoming available in the next month. Therefore, it is possible that our model captures a much larger part of the variation that is indeed predictable. Second, our bid-ask spread forecast is with a standard deviation of 0.27% for the first differences much less noisy compared to realized liquidity with a standard deviation of 0.84%. Therefore, it is possible that the results using our forecasts are less prone to regression attenuation, which biases coefficients towards zero when regressors are measured with errors.

3 Fund Flows and Expected Liquidity Deterioration

In this section, we investigate the impact of expected asset liquidity deterioration on corporate bond fund flows. Investors monitor their mutual funds very closely and reallocate money based on previous performance. However, the shape of this flow-performance relation differs between markets. For equity funds, investors reward funds stronger for good performance and are less sensitive to poor performance (see, e.g., Huang, Wei, and Yan (2007)). In contrast, investors in corporate bond funds are rather insensitive to good past performance, but very sensitive to poor past performance. Goldstein, Jiang, and Ng (2017) argue that the concavity of the flow-performance relation for corporate bond funds is related to strategic complementarities among fund investors and the mismatch between the liquidity a fund offers and the illiquidity of its portfolio holdings. If an investor sells her shares, she usually gets the net asset value as of the time of sale. However, portfolio readjustments happen at a later day and liquidation costs then impose negative externalities on the investors who remain in the fund. These effects are stronger when the fund holds more illiquid bonds. Investors take such a first-mover advantage into account and try to preempt other investors when they are considering redeeming the shares of a poorly performing fund.[20] In aggregate, such a behavior can lead to 'runs' on funds similar to bank runs and impair financial stability (see also Chen, Goldstein, and Jian, 2010).

Ultimately, in such a redemption cascade, the first investors get the best outcomes. For that reason, we expect that investors actively try to anticipate liquidity deterioration of poorly performing funds and act on this expectation.

3.1 Methodology

We test our hypothesis by examining the flow pattern in corporate bond funds when the liquidity of the fund's assets is expected to decrease. To do so, we select an approach which is inspired by Goldstein, Jiang, and Ng (2017) and perform a monthly panel regression of flows in corporate bond funds on expected fund asset liquidity changes:

$$\begin{aligned} \text{Flow}_{k,t} = &\beta_0 + \beta_1 \cdot \text{ExpLiqChange}_{k,t+1} + \beta_2 \cdot \text{ExpLiqChange}_{k,t+1} \cdot \mathbb{1}_{\{\text{Alpha}_{k;t-12,t-1}<0\}} \\ &+ \beta_3 \cdot \text{FundLiq}_{k,t} + \beta_4 \cdot \text{FundLiq}_{k,t} \cdot \mathbb{1}_{\{\text{Alpha}_{k;t-12,t-1}<0\}} \\ &+ \beta_5 \cdot \mathbb{1}_{\{\text{Alpha}_{k;t-12,t-1}<0\}} + \gamma \cdot \text{Controls}_{k,t} + \epsilon_{k,t}, \end{aligned} \tag{4}$$

where $\text{Flow}_{k,t}$ is the flow of fund k in month t and $\mathbb{1}_{\{\text{Alpha}_{k;t-12,t-1}<0\}}$ is a dummy variable that equals 1 if the past performance of fund k is negative and 0 otherwise. Because there is no reason for investors to anticipate a redemption cascade when past performance is positive, we differentiate between funds with positive and negative performance to analyze the relation of expected liquidity changes and fund flows. Note that the anticipated liquidity change for the month ahead $\text{ExpLiqChange}_{k,t+1}$ can be calculated with information from t. Further, to distinguish between investors who trade on anticipations and those who only use realized liquidity, we include the fund's current asset liquidity $\text{FundLiq}_{k,t}$ as well as an interaction term with past performance. We also include the dummy variable for negative past performance separately and expect that flows are negative when this dummy variable equals 1. Lastly, we follow Goldstein, Jiang, and Ng (2017) and control for a fund's flow in

[20] In 2016, the SEC adopted a rule that allows funds to adjust their NAVs to reflect liquidation costs (swing pricing). The rule became effective in November 2018 (see also Capponi, Glasserman, and Weber, 2020; Jin, Kacperczyk, Kahraman, and Suntheim, 2019).

the previous month, total net assets, age, net expense ratio, if redemption fees are charged, and monthly fixed effects.

Monthly flows in individual funds build the basis for our analysis.[21] We calculate the flow for fund k in month t as the relative monthly change in total net assets $\text{TNA}_{k,t}$, adjusted for the fund's return $r_{k,t}$, i.e., $\text{Flow}_{k,t} = \frac{\text{TNA}_{k,t} - \text{TNA}_{k,t-1}(1+r_{k,t})}{\text{TNA}_{k,t-1}}$, using corporate bond fund data from Morningstar (see Appendix B for more details on the data). Following the standard practice in the literature, we winsorize fund flows each month at the 1% and 99% levels. Consistent with Goldstein, Jiang, and Ng (2017), we measure fund liquidity and expected liquidity changes as value-weighted averages across the corporate bonds in the fund's portfolio. Because portfolio holdings are reported at month ends, we use the holdings from the previous month $t-1$ so that the fund flows of the current month cannot influence portfolio compositions, and merge them with asset liquidity in t. For funds that report holdings only quarterly, we use portfolio compositions back to month $t-3$. Liquidity is measured using the average bid-ask spread measure from Section 2.1. For expected bond liquidity, we employ the predictions of our restricted forecasting model of Section 2.3. The expected change in fund k's liquidity for month $t+1$ is then just the relative difference between expected liquidity for $t+1$ and current asset liquidity.

For the separation of funds regarding their past performance, we calculate a fund's average alpha in the preceding twelve months. Again, we follow Goldstein, Jiang, and Ng (2017) and perform time-series regressions of excess fund returns on excess aggregate bond market and stock market returns using a rolling window of the past twelve months. We use the Vanguard Total Bond Market Index Fund return to approximate the aggregate bond market return and the CRSP value-weighted market return for the aggregate stock market return. Fund k's $Alpha_{k;t-12,t-1}$ at month t is then just the estimated intercept of the rolling regression. Finally, for the control variables, we calculate the logarithms of the fund's age and total net assets and create an indicator variable $RearLoad_k$ which equals 1 if the fund charges rear load fees and 0 otherwise.

[21]Goldstein, Jiang, and Ng (2017) argue that fund share-level characteristics such as expense ratios, management fees, and redemption fees can have an influence on investor reallocation decisions and thus use individual fund share classes as unit of observation. We follow them and, for ease of readability, use fund and fund share class as synonym for the rest of the paper.

3.2 Results

The results of the panel regression (4) are reported in Table 5. Before examining the effect of an expected liquidity deterioration, we analyze the relation of flows and the funds' current asset liquidity. To this end, we estimate regression (4) without the expected fund liquidity change. This setting is comparable to the original approach of Goldstein, Jiang, and Ng (2017), which targets the amplification of flows out of poorly performing illiquid funds. The main difference here is that we change perspective and set the focus on a fund's asset liquidity rather than on its performance.[22] The results are given in the first specification of Table 5. Consistent with Goldstein, Jiang, and Ng (2017), we see a significantly negative coefficient for the interaction of fund asset liquidity and fund performance. An increase of 100 basis points of the average bond bid-ask spread leads to an additional outflow of roughly 0.61% for poorly performing funds. We can interpret this result as a 'run' effect out of struggling funds exacerbated by asset illiquidity. For funds with a positive performance over the last year, we find an insignificant effect on their flows. This finding is consistent with the intuition that positive performance convinces most investors to stay in the fund and thus they are not pressured to sell their shares to avoid the negative externalities of others leaving the fund. On the contrary, it is likely that these funds harvest illiquidity premiums, which might contribute to their good performance. As expected, a negative performance as a standalone dummy variable is associated with a significant outflow. Regarding the control variables, we find that corporate bond fund flows show a significantly positive autocorrelation. Also, older funds experience relatively more outflows. While the fund's total net assets and its net expense ratio do not have a significant impact on its flows, charging rear load fees is, consistent with Goldstein, Jiang, and Ng (2017), significantly associated with flows out of the fund.

Because Goldstein, Jiang, and Ng (2017) show that the run effect is stronger for worse performing funds, we examine three further specifications in which we employ alternative cutoffs for the alpha indicator variable. We use cutoffs that correspond to the upper quartile, the median, and the lower quartile of all negative (monthly) alphas, which are -5 bps, -12.5 bps, and -30 bps, respectively. Across the four specifications (1), (3), (5), and (7) of Table 5, we see that the amplification effect is indeed stronger for worse performing funds. While a 100 basis points increase in the average bid-ask spread is associated with

[22] Goldstein, Jiang, and Ng (2017) also examine the interaction between performance and fund illiquidity for funds with negative past alpha, employing performance as continuous variable and fund liquidity as indicator variable. In the robustness section 4.3, we include performance measured via a continuous variable as an additional control.

a cumulated outflow of 0.53% (= 0.0747 − 0.6082, see third-to-last line of Table 5) for all negative performing funds in specification (1), this outflow increases to 1.10% when the fund's alpha belongs to the 25% most negative ones (specification (7)). Thus, the cumulated liquidity effect for the lowest alpha quartile is more than doubled compared to the effect for all funds with negative performance.

Insert Table 5 about here.

We now present the results on our main hypothesis that investors anticipate a liquidity deterioration and exit the fund based on their expectation using the full regression model (4). In specifications (2), (4), (6), and (8), we always control for the current fund liquidity. This variable captures the costs that investors expect the fund to incur when it has to liquidate assets in response to outflows. Therefore, the effect of an expected liquidity change can be interpreted as investors' attempt to preempt other investors based on expected changes of these costs in the future. While specification (2) of Table 5 is based on the indicator variable that simply separates funds according to the sign of their alpha, specifications (4), (6), and (8) use the three different negative performance cutoffs from above. The interaction term between the expected liquidity change and fund performance is negatively significant with t-statistics between 4 and 5 in all specifications. Looking at the cumulated effect, we find that an expected doubling of the bid-ask spread (i.e., an increase of 100%) is associated with significantly stronger outflows of 0.69% for funds with a negative alpha over the last year. As hypothesized, investors seem to leave poorly performing funds in advance if a deterioration in liquidity is expected. For funds with a positive alpha, we see that anticipated changes in liquidity have no effect on flows, which is consistent with the intuition that their investors are not pressured to act strategically. The cumulated anticipation effect is in all four settings statistically significant. Most importantly, because the incentive for investors to sell their shares as soon as possible increases for worse performing funds, the effect increases monotonously when the performance becomes worse. While an expected 100% increase of the bid-ask spread is associated with an outflow of roughly 0.69% for all funds with a negative alpha, this response increases to 0.90% for the 50% worst performing funds and more than doubles to 1.46% for the 25% worst funds. Consistent with our previous results, we find in each setting a significantly higher flow out of poorly performing funds if their assets are currently more illiquid. Further, we see that the cumulated effect is always statistically significant and again becomes stronger for worse performing funds.

Next, we want to compare the magnitude of the anticipation effect with the effect of realized liquidity. To this end, we calculate the economic significance of both effects for the four different settings at the bottom of Table 5. On the one hand, a one standard deviation decrease in (current) fund liquidity is associated with an outflow of 0.28% for funds with negative alpha. This effect increases to 0.65% when the fund's alpha is below −30 bps and thus belongs to the 25% worst funds with negative performance. On the other hand, a one standard deviation of expected liquidity deterioration leads to an outflow of 0.16% for negatively performing funds, which increases to 0.45% for the worst quarter of poor funds. This finding again emphasizes that with decreasing performance, the pressure on investors to redeem their shares immediately is stronger. Comparing both effects, the anticipation effect is roughly 60% to 70% of the size of the realized liquidity effect across the four settings. The comparison with the settings in which we exclude liquidity expectations (specifications (1), (3), (5), and (7)) shows that the economic significance of realized liquidity even slightly increases when including expected liquidity changes. This finding emphasizes that both effects indeed represent separate channels contributing to liquidity-induced 'fund runs'. Thus, only looking at realized liquidity severely underestimates the magnitude of these 'runs' compared to the more complete picture that we provide.

Finally, we want to shed more light on the mechanism behind the anticipation effect. Since strategic complementarities arise for investors only in case of liquidity deterioration, we hypothesize that expected changes in fund liquidity are related asymmetrically to fund flows. To test this hypothesis, we split the anticipated liquidity change $\text{ExpLiqChange}_{k,t+1}$ in equation (4) into anticipated liquidity deterioration and improvement. The results are presented in Table 6. For poorly performing funds, we observe indeed an asymmetric pattern that only expected liquidity deteriorations have a significant effect on fund flows. Lastly consistent with our previous findings, we see that both expected liquidity deterioration and improvement do not have a significant effect on well performing funds.

Insert Table 6 about here.

Summarizing, our results are consistent with investors actively anticipating liquidity deterioration in underperforming funds. Such forecasts will provide them with a first-mover advantage when selling their shares in advance. From the perspective of financial stability, this behavior is dangerous as it could trigger redemption spirals.

4 Robustness

In this section, we show that our results are robust against three critical alternative specifications. We first compare our linear forecasting model with a random forest model and show that the empirical results of Sections 2.5 and 3 are robust also for the random forest. Second, we use a more sophisticated liquidity measure that takes the size dependence of transaction costs into account. Third, we rerun the analyses in Section 3 using additional variables and interactions for the funds' previous performance.

4.1 Random Forest Prediction Model

In machine learning, the random forest model has become popular as a rather simple non-parametric alternative to classic linear prediction models (see, e.g., Behrens, Pierdzioch, and Risse, 2018; Gu, Kelly, and Xiu, 2020). The basis for this algorithm are regression trees that try to find similar groups among the observations. Simply speaking, the algorithm adds branches to the tree at each step by sorting the observations of the previous node into these new branches based on one of the predictor variables. At the end of this procedure, the terminal nodes (leafs) form the partitions and an observation's predicted value is then the average of the left-hand side variable from all observations in the same partition. Since the predictor variables used for branching as well as their cutoff are chosen to get the minimal forecasting error on the validation set, a single regression tree is prone to overfitting. To overcome this limitation, the random forest model employs a bagging procedure that fits a regression tree for n randomly drawn samples. The final prediction is then simply the average across the predictions of the n regression trees.[23] We employ the random forest procedure with $n = 500$ to our prediction approach of Section 2.3 for both the full and the restricted set of candidate predictors. Similar to the out-of-sample cross-validation methods in Section 2.3, we use a rolling window to fit the model. In each month t, we split the observations of the previous twelve months into a training set of eleven months and a validation set of the last month.

The predictive accuracy of the random forest model compared to the naïve benchmark is shown in Panel A of Table 7. For both the full and the restricted model, we see a decrease in RMMSE of roughly 20%, which is very similar to the improvement of the linear forecasting model in Panel A of Table 3. The relation between yield spreads and expected liquidity also

[23] We refer for a more detailed description of the random forest algorithm to Gu, Kelly, and Xiu (2020).

remains virtually unchanged if we use the random forest instead of our linear model as a forecasting method (see Panel A of Table 8). If anything, the effect of expected liquidity is stronger.

Insert Table 7 and Table 8 about here.

We also analyze the robustness of our results from Section 3 that investors consider realized liquidity and anticipated liquidity deteriorations to secure a first-mover advantage when redeeming their corporate bond fund shares. In Table 9, we employ the random forest model to calculate expected liquidity changes. Consistent with our previous findings, we see in specifications (2), (4), (6), and (8) highly significant interaction terms with both realized liquidity and expected liquidity changes. Again, both effects increase in the economic significance from all funds with negative performance (specification (2)) to only funds with the most negative alphas (specification (8)). Quantitatively, the effects are slightly weaker compared to the linear model in Table 5.

Insert Table 9 about here.

4.2 Size-Adapted Liquidity Measure

Transaction costs in bond markets strongly depend on trade size (see, e.g., Edwards, Harris, and Piwowar, 2007). For that reason, idiosyncratic or systematic variations in the trade size disturb measured transaction costs. Such a bias is particularly strong for those bonds that trade only a few times in a month. In a recent paper, Reichenbacher and Schuster (2020) develop a procedure to mitigate the related measurement errors. Their basic idea is to compare the paid transaction costs for a given volume to the costs usually paid for similar trade sizes. The trade-size adapted liquidity measure is then given as the scaling factor between the observed and the usually paid costs. To rule out that our results are influenced by such measurement errors, we follow them and employ daily average bid-ask spreads to estimate monthly scaling factors $sf_{i,t}$ using the following model

$$AvgBidAsk_{i,d} = sf_{i,t} \cdot \frac{1}{2}\left[\frac{1}{n_{i,d}^{buy}}\sum_{k=1}^{n_{i,d}^{buy}} c(vol_{k,i,d}^{buy}) + \frac{1}{n_{i,d}^{sell}}\sum_{k=1}^{n_{i,d}^{sell}} c(vol_{k,i,d}^{sell})\right] + \epsilon_{i,d},$$

where $c(vol)$ characterizes the usual relation between transaction costs and trade size and $n_{i,d}^{buy/sell}$ gives the number of buy and sell trades in bond i on day d.[24] Note that the scaling factor is a relative measure of liquidity. For example, a scaling factor of $sf_{i,t} = 2$ means that bid-ask spreads for bond i in month t are twice as large compared to the average bond in the sample. Therefore, we can use the scaling factors instead of the standard average bid-ask spread measure and apply our forecasting methodology of Section 2.3.

The results regarding the predictive accuracy of the forecasting model and the naïve benchmark are presented in Panel B of Table 7. Note that the naïve benchmark now assumes that the size-adapted liquidity measure, i.e., the scaling factor is unchanged compared to the previous month. We find a reduction of the RMMSE for this liquidity measure of about 22% for the full and the restricted model compared to the naïve benchmark. The results of our indirect performance evaluation that relies on the relation between changes in yield spreads and changes in expected liquidity are shown in Panel B of Table 8. Consistent with the findings in Reichenbacher and Schuster (2020), we find that the modified measure in specification (2) based on the naïve forecast leads to a higher explanatory power and lower out-of-sample error compared to specification (2) in Table 4. The explanatory power and out-of-sample error are further improved strongly when moving to differences in expected liquidity calculated with our linear forecasting model in specification (3). The improvements in the R^2_{adj} and the out-of-sample MSE are more than twice as large compared to Table 4. One explanation for the stronger improvement is that this measure does not suffer from noise in the estimation due to changing trade-size patterns and as a result can be predicted more precisely. In summary, we also find for this measure that our forward-looking approach directly and indirectly outperforms the naïve benchmark in terms of predictive accuracy.

Next, we employ the size-adapted average bid-ask spread in our fund flow analysis of Section 3. The results are provided in Panel B of Table 9. For all performance quantiles, we consistently find significantly negative interaction effects of realized liquidity and expected liquidity changes. The impact of both effects on corporate bond fund flows becomes again stronger with decreasing fund performance.

4.3 Fund Flows and Alpha

Goldstein, Jiang, and Ng (2017) find that outflows are stronger for more negative alphas. In our main analysis, we capture this dependence with dummy variables for different thresholds

[24] For further details on the size-adapted measure and its implementation, we refer to Reichenbacher and Schuster (2020).

of alpha. In Panel C of Table 9, we additionally include alpha as a continuous variable and an interaction term with the dummy variable for negative performance. Note that we have to use alpha minus the alpha-threshold in the interaction term so that the functional relation between alpha and flows is continuous. Looking at specification (1), where the dummy variable is one if alpha is negative, we find both a significantly positive coefficient for the continuous alpha and a significantly positive interaction term. This effect is consistent with Goldstein, Jiang, and Ng (2017) and confirms the concavity of fund flows. A response of flows to performance becomes stronger when the alpha is negative. Interestingly, for more negative thresholds, the interaction variable first becomes insignificant and for the most extreme setting, where the dummy variable is only one for the 25% most negative alphas, becomes negative (note that alpha plus the interaction term is still positive). This result indicates that for strongly negative alphas, the relation to flows becomes again flatter.

Most importantly, our results regarding the relation of observed liquidity and expected liquidity changes with the outflows of poorly performing funds remain robust. The cumulated effects and the economic significance again increase with decreasing fund performance.

5 Conclusion

In this study, we propose a prediction model for individual bond liquidity. Our methodology incorporates a dynamic predictor selection that, first, accounts for the information available to a contemporary forecaster and, second, allows for changing structural relations between the economic variables over time. Our approach is easy to implement for any liquidity measure and we exemplarily apply it to the average bid-ask spread measure.

To assess the performance of the forward-looking approach, we compare it to the literature's current practice to use today's liquidity as the best predictor of future liquidity. We show that our forecasting model outperforms this naïve benchmark in every month of our sample period. The method performs also remarkably well during the financial crisis. Additionally, we perform an indirect performance test exploiting the relation of changes in market expectations for future liquidity and yield-spread changes. The new prediction model reveals a much stronger dependence of yield spreads on liquidity changes and strongly increases the explainable part of yield spread changes.

Finally, we use our predictions to examine the impact of declining liquidity in poorly performing corporate bond funds on investors' selling decision. Consistent with the implications of strategic complementarities among corporate bond fund investors, we find two types of investor behavior. First, investors actively anticipate liquidity deteriorations in funds with poor past performance and sell their fund shares in advance to secure a first-mover advantage. Second, investors also react stronger for funds with already illiquid portfolio holdings. Both effects become more pronounced for funds with more negative performance during their last twelve months. Our results emphasis the importance of the recent regulatory change that allows to pass on the costs of redemptions to the redeeming shareholders (swing pricing). Because the fire sales from struggling funds might also impact the market as a whole, the regulator should incentivize funds to broadly use this new possibility (for a similar argument, see the theoretical model in Capponi, Glasserman, and Weber, 2020).

References

Acharya, V. V., Y. Amihud, and S. T. Bharath, 2013, Liquidity Risk of Corporate Bond Returns: Conditional Approach, *Journal of Financial Economics* 110, 358–386.

Agarwal, V. and N. Naik, 2004, Risks and Portfolio Decisions Involving Hedge Funds, *The Review of Financial Studies* 17, 63–98.

Amihud, Y., 2002, Illiquidity and Stock Returns: Cross-Section and Time Series Effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Y. and H. Mendelson, 1986, Asset Pricing and the Bid-Ask Spread, *Journal of Financial Economics* 17, 223–249.

Bao, J., J. Pan, and J. Wang, 2011, The Illiquidity of Corporate Bonds, *Journal of Finance* 66, 911–946.

Behrens, C., C. Pierdzioch, and M. Risse, 2018, A Test of the Joint Efficiency of Macroeconomic Forecasts using Multivariate Random Forests, *Journal of Forecasting* 37, 560–572.

Bongaerts, D., F. de Jong, and J. Driessen, 2017, An Asset Pricing Approach to Liquidity Effects in Corporate Bond Markets, *Review of Financial Studies* 30, 1229–1269.

Boyarchenko, Nina, Domenico Giannone, and Or Shachar, 2019, Flighty liquidity, Working paper.

Capponi, Agostino, Paul Glasserman, and Marko Weber, 2020, Swing pricing for mutual funds: Breaking the feedback loop between fire sales and fund redemptions, *Management Science*.

Chen, Q., I. Goldstein, and W. Jian, 2010, Payoff Complementarities, Self-fulfilling Beliefs and Mutual Fund Outflows, *Journal of Financial Economics* 98, 239–262.

Chernobai, A., P. Jorion, and F. Yu, 2011, The Determinants of Operational Risk in US Financial Institutions, *Journal of Financial and Quantitative Analysis* 46, 1683–1725.

Chinco, A., A. Clark-Joseph, and M.o Ye, 2019, Sparse Signals in the Cross-Section of Returns, *The Journal of Finance* 74, 449–492.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of financial economics* 56, 3–28.

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, An Empirical Analysis of Stock and Bond Market Liquidity, *The Review of Financial Studies* 18, 85–129.

Dick-Nielsen, J., 2009, Liquidity Biases in TRACE, *Journal of Fixed Income* 19, 43–55.

Dick-Nielsen, J., 2014, How to Clean Enhanced TRACE Data, Working paper.

Dick-Nielsen, J., P. Feldhütter, and D. Lando, 2012, Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis, *Journal of Financial Economics* 103, 471–492.

Diebold, F. and R. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business & Economic Statistics* 13, 253–263.

Downing, C., S. Underwood, and Y. Xing, 2005, Is liquidity Risk Priced in the Corporate Bond Market?, Working paper.

Duffee, G. R., 1998, The Relation between Treasury Yields and Corporate Bond Yield Spreads, *The Journal of Finance* 53, 2225–2241.

Edwards, A. K., L. E. Harris, and M. S. Piwowar, 2007, Corporate Bond Market Transaction Costs and Transparency, *Journal of Finance* 62, 1421–1451.

Fama, E. and J. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of political economy* 81, 607–636.

Friewald, N., R. Jankowitsch, and M. G. Subrahmanyam, 2012, Illiquidity or Credit Deterioration: A Study of Liquidity in the US Corporate Bond Market During Financial Crisis, *of Financial Economics* 105, 18–36.

Gehde-Trapp, M., P. Schuster, and M. Uhrig-Homburg, 2018, The Term Structure of Bond Liquidity, *Journal of Financial and Quantitative Analysis* 53, 2161–2197.

Giacomini, R. and H. White, 2006, Tests of Conditional Predictive Ability, *Econometrica* 74, 1545–1578.

Goldstein, I., H. Jiang, and D. T. Ng, 2017, Investor Flows and Fragility in Corporate Bond Funds, *Journal of Financial Economics* 126, 592–613.

Goyenko, R. and A. Ukhov, 2009, Stock and Bond Market Liquidity: A Long-Run Empirical Analysis, *Journal of Financial and Quantitative Analysis* 44, 189–212.

Gu, S., B. Kelly, and D. Xiu, 2020, Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies* 33, 2223–2273.

Gürkaynak, R., B. Sack, and J. Wright, 2007, The US Treasury Yield Curve: 1961 to the Present, *Journal of Monetary Economics* 54, 2291–2304.

Harvey, D., S. Leybourne, and P. Newbold, 1997, Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting* 13, 281–291.

He, Z. and K. Milbradt, 2014, Endogenous Liquidity and Defaultable Bonds, *Econometrica* 82, 1443–1508.

He, Z. and W. Xiong, 2012, Rollover Risk and Credit Risk, *The Journal of Finance* 67, 391–430.

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

Hotchkiss, E. and G. Jostova, 2017, Determinants of Corporate Bond Trading: A Comprehensive Analysis, *Quarterly Journal of Finance* 7, 1750003.

Huang, J., K. Wei, and H. Yan, 2007, Participation Costs and the Sensitivity of Fund Flows to Past Performance, *The Journal of Finance* 62, 1273–1311.

Jankowitsch, R., A. Nashikkar, and M. Subrahmanyam, 2011, Priced Dispersion in OTC Markets: A New Measure of Liquidity, *Journal of Banking and Finance* 35, 343–357.

Jin, Dunhong, Marcin T Kacperczyk, Bige Kahraman, and Felix Suntheim, 2019, Swing Pricing and Fragility in Open-End Mutual Funds, .

Kacperczyk, M., S. Nieuwerburgh, and L. Veldkamp, 2014, Time-varying Fund Manager Skill, *The Journal of Finance* 69, 1455–1484.

Kozak, S., S. Nagel, and S. Santosh, 2020, Shrinking the Cross-Section, *Journal of Financial Economics* 135, 271–292.

Liu, X. and J. R. Ritter, 2011, Local Underwriter Oligopolies and IPO Underpricing, *Journal of Financial Economics* 102, 579–601.

Mahanti, S., A. Nashikkar, M. Subrahmanyam, G. Chacko, and G. Mallik, 2008, Latent Liquidity: A New Measure of Liquidity, with an Application to Corporate Bonds, *Journal of Financial Economics* 88, 272–298.

Panopoulou, E. and S. Vrontos, 2015, Hedge Fund Return Predictability; To Combine Forecasts or Combine Information?, *Journal of Banking & Finance* 56, 103–122.

Pesaran, M. and A. Timmermann, 1995, Predictability of Stock Returns: Robustness and Economic Significance, *The Journal of Finance* 50, 1201–1228.

Rapach, D., J. Strauss, and G. Zhou, 2010, Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *The Review of Financial Studies* 23, 821–862.

Reichenbacher, M. and P. Schuster, 2020, Size-Adapted Bond Liquidity Measures and Their Asset Pricing Implications, Working paper.

Schestag, R., P. Schuster, and M. Uhrig-Homburg, 2016, Measuring Liquidity in Bond Markets, *Review of Financial Studies* 29, 1170–1219.

Smith, J. and K. Wallis, 2009, A Simple Explanation of the Forecast Combination Puzzle, *Oxford Bulletin of Economics and Statistics* 71, 331–355.

Titman, S. and C.n Tiu, 2011, Do the Best Hedge Funds Hedge?, *The Review of Financial Studies* 24, 123–168.

Warga, Arthur, 1992, Bond returns, liquidity, and missing data, *Journal of Financial and Quantitative Analysis* 27, 605–617.

A Bond Data Filters and Yield Spread Calculation

We use transaction data from Enhanced TRACE for the U.S. corporate bond market for the time period from October 1, 2004 to June 30, 2017. Despite TRACE being gradually introduced starting in 2002 by the FINRA, we restrict our examination period to a start in October 2004. After this date, market participants had to report almost all trades. To clean the data, we apply several filters comparable to the ones used in Schestag, Schuster, and Uhrig-Homburg (2016) and Reichenbacher and Schuster (2020). In a first step, we apply the standard procedures of Dick-Nielsen (2009, 2014) to remove duplicates, withdrawn and corrected entries. In the second step, we apply the median and reversal filter of Edwards, Harris, and Piwowar (2007) to exclude erroneous entries and extreme outliers. Finally, we demand bonds to satisfy several conditions. First, we exclude bonds with special features (perpetuals, convertible and puttable bonds, floating coupon payments). Second, bonds have to be senior unsecured, USD denominated, and there shall be no entity backing the bond with a guarantee. Third, we demand bonds to be actively traded in at least twelve months within our sample period or in 50% of the months they are active. Last, we exclude observations of defaulted bonds after the default happened. Overall, our sample consists of 25,918 bonds and 61,360,046 trades, which is approximately 57% of the about 108 million trades during our observation period.

We require a bond's yield spread for the analysis in Section 2.5. If the yield-to-maturity is larger than the yield-to-call, the reported yield in TRACE often (but not always) corresponds to the yield-to-call. We calculate both yields and select the one that is closest to the reported yield. We drop observations for which both differences are larger than 1 basis point (about 0.7% of all observations). We discount the bond's cash flows with the risk-free Treasury curve to calculate the price of an artificial Treasury bond with the same cash flow structure (using updated data from Gürkaynak, Sack, and Wright (2007) published by the Federal Reserve on http://www.federalreserve.gov). Finally, the yield spread is defined as the continuously compounded yield computed from the reported price minus the corresponding yield calculated from the price of the artificial Treasury bond (see also Gehde-Trapp, Schuster, and Uhrig-Homburg, 2018).

B Mutual Fund Data

We use mutual fund data from Morningstar. We identify corporate bond funds using the Morningstar category classifications and the fund's prospectus objective. Specifically, funds have to be in one of the categories 'US Fund Long-Term Bond', 'US Fund Intermediate Core Bond', 'US Fund Intermediate Core-Plus Bond', 'US Fund Short-Term Bond', 'US Fund Ultrashort Bond', 'US Fund Corporate Bond', 'US Fund High Yield Bond', 'US Fund Target Maturity', or 'US Fund Multisector Bond'. Additionally, we exclude funds with a focus on government or municipal bonds in their prospectus objective. The remaining filters follow Goldstein, Jiang, and Ng (2017). We exclude index funds as well as fund share classes during their first year. Finally, to get sufficient data coverage when merging fund holding data with TRACE, we restrict the sample in our analyses to the time period from January 2008 to June 2017. After applying these filters, our sample consists of 3,492 share classes from 1,005 corporate bond funds.

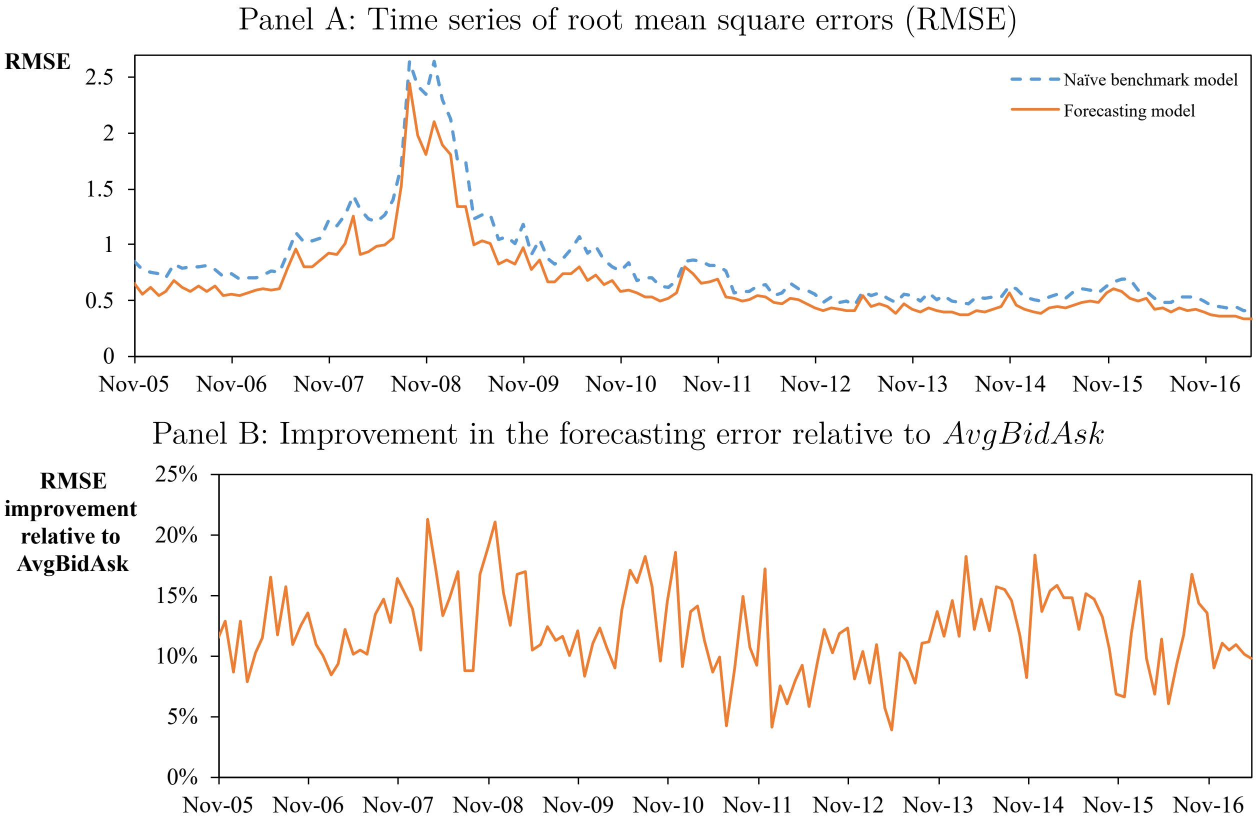

Figure 1: **Forecasting performance**

This figure shows the time series performance of the naïve benchmark model and the linear combination model (fc) of Section 2.3 for the average bid-ask spread measure $AvgBidAsk$. In Panel A, we report the time series of the monthly root mean square errors (RMSE). The blue (dashed) line represents errors of the naïve benchmark model and the orange (solid) line the errors of the forecasting model. Panel B shows the improvement in the forecasting error of the forecasting model relative to the average liquidity, i.e., $\frac{RMSE_t^{naïve} - RMSE_t^{fc}}{AvgBidAsk_t}$.

Table 1: **Descriptive statistics**
This table shows descriptive statistics for the set of candidate predictors (see Section 2.2). In Panel A, we report average cross-sectional statistics for transaction cost variables, bond characteristics, trading activity, and credit quality variables that have both time-series and cross-sectional variation. To this end, we first calculate each month the cross-sectional statistics and then average over time. In Panel B, we report time series statistics for the predictor variables on the bond market, the equity market, and the broader economic environment. We measure bond liquidity using the average bid-ask spread measure.

Panel A: Variables with time-series and cross-sectional variation							
	Mean	Std. dev.	$Q_{5\%}$	$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$	$Q_{95\%}$
Current liquidity (%)	1.30	1.22	0.05	0.42	0.96	1.87	3.65
12-month mov. avg. Liquidity (%)	1.32	1.02	0.19	0.54	1.06	1.89	3.29
Segment liquidity (%)	1.29	0.67	0.44	0.81	1.13	1.67	2.65
Volatility of liquidity (%)	0.49	0.61	0.00	0.01	0.32	0.72	1.67
Bond return (%)	0.48	2.88	-3.35	-0.59	0.39	1.49	4.52
Age (years)	4.30	3.71	0.43	1.73	3.43	5.77	11.81
Duration (years)	6.07	4.64	0.57	2.47	4.87	8.34	15.51
Amount outstanding (log USD)	19.19	1.65	15.84	18.23	19.64	20.35	21.23
Average trade size (log USD)	12.36	1.51	9.65	11.20	12.74	13.53	14.27
Total trading volume (log USD)	17.63	2.26	13.41	16.10	18.16	19.31	20.56
Bond order imbalance (%)	-0.61	35.05	-65.43	-16.95	-0.41	16.14	62.43
Rating (1: AAA,..., 22: D)	7.75	3.40	2.42	5.52	7.38	9.74	14.19
CDS spread (%)	1.73	3.14	0.28	0.47	0.84	1.68	6.12
Panel B: Variables with time-series variation							
	Mean	Std. dev.	$Q_{5\%}$	$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$	$Q_{95\%}$
Bond market liquidity (%)	1.29	0.48	0.73	0.86	1.29	1.55	2.24
1-month bond market return (%)	0.43	1.62	-1.84	-0.45	0.45	1.37	2.77
12-month bond market return (%)	5.44	6.43	-4.26	1.65	5.14	8.15	18.68
Bond market order imbalance (%)	-0.64	2.29	-4.28	-1.90	-0.60	0.86	2.54
Bond market volatility (%)	4.79	1.80	2.75	3.52	4.34	5.72	8.31
TYVIX (%)	6.30	1.93	4.30	4.92	5.66	7.03	10.49
Stock market liquidity	0.042	0.018	0.022	0.028	0.040	0.050	0.074
1-month stock market return (%)	0.76	4.02	-7.03	-1.49	1.27	3.27	6.78
12-month stock market return (%)	9.38	15.99	-32.57	4.91	12.11	17.37	30.20
VIX (%)	19.00	8.98	11.26	13.43	16.19	21.61	36.53
Inflation (%)	2.06	1.47	-0.20	1.15	1.99	3.17	4.31
1-month TED spread (%)	0.38	0.46	0.12	0.16	0.20	0.36	1.40
6-month TED spread (%)	0.52	0.39	0.23	0.28	0.35	0.60	1.31
Federal funds rate (%)	1.36	1.86	0.08	0.12	0.19	2.39	5.25
Industrial production growth (%)	0.69	4.70	-11.63	-0.84	2.26	3.25	5.31

Table 2: **Correlation matrix**

This table reports correlations for next month's bond liquidity and the candidate predictors (see Section 2.2) based on panel data. We measure bond liquidity using the average bid-ask spread measure.

	Next month's liquidity	Current liquidity	12-month liquidity	Segment liquidity	Volatility of liquidity	Bond return	Age	Duration	Amount outstanding	Average trade size	Total trading volume	Bond order imbalance	Rating	CDS spread	Bond market liquidity
Next month's liquidity	1	0.71	0.73	0.63	0.40	0.00	0.17	0.32	-0.43	-0.48	-0.31	-0.01	0.09	0.27	0.36
Current liquidity		1	0.75	0.65	0.47	0.02	0.17	0.31	-0.43	-0.50	-0.32	-0.03	0.09	0.28	0.37
12-month liquidity			1	0.69	0.39	0.07	0.24	0.34	-0.54	-0.60	-0.40	0.00	0.10	0.28	0.37
Segment liquidity				1	0.32	0.05	0.10	0.41	-0.62	-0.55	-0.50	-0.02	0.11	0.23	0.56
Volatility of liquidity					1	0.03	0.08	0.20	-0.04	-0.11	0.04	-0.01	0.19	0.28	0.28
Bond return						1	0.02	0.04	-0.02	0.00	0.00	-0.05	0.04	0.00	0.05
Age							1	-0.02	-0.12	-0.13	0.00	0.01	0.09	0.09	0.00
Duration								1	-0.08	-0.06	-0.03	-0.01	-0.07	-0.04	-0.02
Amount outstanding									1	0.81	0.89	0.00	0.00	-0.12	-0.24
Average trade size										1	0.80	0.01	0.15	0.01	-0.22
Total trading volume											1	0.01	0.10	0.06	-0.17
Bond order imbalance												1	0.01	0.00	-0.01
Rating													1	0.50	-0.05
CDS spread														1	0.20
Bond market liquidity															1
Bond market return (1m)															
Bond market return (12m)															
Bond market order imbalance															
Bond market volatility															
TYVIX															
Stock market liquidity															
Stock market return (1m)															
Stock market return (12m)															
VIX															
Inflation															
TED spread (1m)															
TED spread (6m)															
Federal funds rate															
Industrial production growth															

Table 2 continued

	Bond market return (1m)	Bond market return (12m)	Bond market order imbalance	Bond market volatility	TYVIX	Stock market liquidity	Stock market return (1m)	Stock market return (12m)	VIX	Inflation	TED spread (1m)	TED spread (6m)	Federal funds rate	Industrial production growth
Next month's liquidity	0.01	-0.02	-0.05	0.19	0.25	0.33	-0.08	-0.20	0.24	0.11	0.21	0.21	0.14	-0.14
Current liquidity	0.03	-0.01	-0.06	0.19	0.26	0.34	-0.06	-0.21	0.25	0.09	0.20	0.21	0.13	-0.15
12-month liquidity	0.07	0.15	-0.03	0.15	0.22	0.33	0.00	-0.12	0.19	0.04	0.10	0.12	0.12	-0.15
Segment liquidity	0.05	-0.02	-0.08	0.30	0.40	0.52	-0.10	-0.32	0.40	0.13	0.30	0.33	0.18	-0.25
Volatility of liquidity	0.02	-0.04	-0.03	0.17	0.22	0.27	-0.05	-0.20	0.22	0.04	0.15	0.18	0.06	-0.16
Bond return	0.37	0.10	-0.06	-0.07	0.02	0.08	0.18	-0.04	0.03	-0.12	-0.11	-0.01	-0.04	-0.09
Age	0.01	0.03	0.00	0.03	0.03	-0.01	0.00	0.00	0.04	-0.03	-0.01	0.03	-0.06	-0.02
Duration	0.01	0.02	0.00	0.01	0.01	-0.02	0.00	0.00	0.01	-0.03	-0.01	0.01	-0.04	-0.01
Amount outstanding	-0.01	-0.06	0.04	-0.02	-0.06	-0.21	0.01	0.02	-0.04	-0.15	-0.10	0.00	-0.23	-0.03
Average trade size	-0.01	-0.05	0.04	-0.02	-0.06	-0.20	0.01	0.02	-0.04	-0.15	-0.10	-0.01	-0.22	-0.04
Total trading volume	0.00	-0.03	0.03	0.01	-0.01	-0.16	0.01	0.02	0.01	-0.13	-0.08	0.01	-0.23	-0.03
Bond order imbalance	-0.02	0.01	0.05	-0.01	-0.01	-0.02	-0.01	0.01	-0.01	0.00	-0.01	-0.02	0.00	0.01
Rating	0.00	0.00	0.02	-0.05	-0.06	-0.05	0.02	0.03	-0.06	-0.02	-0.04	-0.05	0.00	0.00
CDS spread	0.03	-0.06	-0.03	0.16	0.20	0.21	-0.05	-0.19	0.22	0.00	0.12	0.21	-0.04	-0.15
Bond market liquidity	0.10	-0.02	-0.15	0.53	0.72	0.93	-0.17	-0.57	0.70	0.22	0.52	0.58	0.33	-0.43
Bond market return (1m)	1	0.20	-0.10	-0.16	0.05	0.15	0.28	-0.09	0.10	-0.20	-0.24	-0.03	-0.09	-0.14
Bond market return (12m)		1	0.07	-0.14	-0.13	-0.09	0.20	0.47	-0.14	-0.16	-0.32	-0.32	-0.20	0.17
Bond market order imbalance			1	-0.14	-0.14	-0.13	0.01	0.10	-0.13	-0.08	-0.14	-0.12	-0.07	-0.03
Bond market volatility				1	0.83	0.48	-0.26	-0.52	0.71	-0.11	0.38	0.57	-0.21	-0.40
TYVIX					1	0.66	-0.23	-0.62	0.84	-0.07	0.42	0.67	-0.21	-0.50
Stock market liquidity						1	-0.06	-0.64	0.69	0.14	0.37	0.52	0.25	-0.45
Stock market return (1m)							1	0.23	-0.33	-0.22	-0.31	-0.25	-0.06	0.02
Stock market return (12m)								1	-0.63	0.12	-0.38	-0.68	-0.03	0.64
VIX									1	-0.03	0.42	0.70	-0.22	-0.47
Inflation										1	0.46	0.14	0.52	0.48
TED spread (1m)											1	0.75	0.41	-0.12
TED spread (6m)												1	0.00	-0.48
Federal funds rate													1	0.19
Industrial production growth														1

Table 3: **Forecasting model**

Panel A of this table reports root mean mean square errors (RMMSE) for the forecasting model and the naïve benchmark. The forecasting model is described in Section 2.3. The naïve forecast for a bond's liquidity in the next month equals the realization of today. We measure bond liquidity using the average bid-ask spread measure. The full model includes all candidate predictors from Section 2.2, whereas the restricted model only includes variables that are directly related to liquidity (see Section 2.4). Panel B shows statistics for the variables that are selected in the forecasting approach for the full model. Variables are ranked by the percentage of months for which they are included. Additionally, we report the percentage of months the variables have a positive and negative coefficient $\widehat{\beta}$ given that they are included and their economic significance (in bps). The economic significance is calculated as the average of the monthly product of the parameter estimate and the variable's standard deviation within the 12-month rolling window used for the model calibration. For the calculation of economic significance, variables that are not selected have a coefficient of 0. All statistics are based on the average parameter across the three individual selection approaches that form the combination model.

Panel A: Forecast performance

	Full model		Restricted model	
	RMMSE	Δ to naïve	RMMSE	Δ to naïve
Forecasting model	0.74	-18.71%	0.80	-18.81%
Naïve benchmark model	0.91		0.99	
Observations	230,790		511,465	

Panel B: Selected variables

Rank	Effect	% included	% positive	% negative	Econ. significance
1	12-month liquidity	99.8%	100.0%	0.0%	41.8
2	Current liquidity	99.8%	100.0%	0.0%	30.4
3	Duration	92.9%	100.0%	0.0%	6.7
4	Average trade size	89.8%	0.0%	100.0%	-5.3
5	CDS spread	80.7%	99.3%	0.7%	3.9
6	Segment liquidity	76.7%	98.5%	1.5%	5.1
7	Total trading volume	72.4%	100.0%	0.0%	4.5
8	Volatility of liquidity	71.7%	100.0%	0.0%	2.7
9	Amount outstanding	65.0%	1.4%	98.6%	-3.4
10	Bond return	57.6%	14.2%	85.8%	-1.5
11	Stock market return (1m)	45.7%	17.1%	82.9%	-1.4
12	Age	42.9%	43.5%	56.5%	-0.2
13	Rating	36.0%	31.1%	68.9%	-0.7
14	Bond market return (1m)	33.3%	70.1%	29.9%	0.4
15	Bond market volatility	32.9%	52.6%	47.4%	0.5
16	Bond market return (12m)	31.9%	12.6%	87.4%	-1.2
17	Bond market order imbalance	31.2%	45.2%	54.8%	-0.3
18	Inflation	28.6%	72.6%	27.4%	0.9
19	Industrial production growth	27.9%	37.0%	63.0%	0.2
20	TED spread (1m)	27.1%	28.6%	71.4%	-0.4
21	Bond order imbalance	25.7%	20.5%	79.5%	-0.2
22	Stock market return (12m)	25.2%	48.8%	51.3%	-0.1
23	TYVIX	24.3%	68.9%	31.1%	0.4
24	VIX	21.9%	69.6%	30.4%	0.0
25	Stock market liquidity	21.2%	67.2%	32.8%	0.2
26	TED spread (6m)	19.8%	63.9%	36.1%	0.4
27	Bond market liquidity	19.5%	69.7%	30.3%	0.4
28	Federal funds rate	18.3%	38.7%	61.3%	0.0

Table 4: **Yield spread regression**

This table reports results of the panel regressions of yield spread changes on changes in expected bid-ask spreads

$$\Delta(\text{Yield spread})_{i,t} = \alpha + \beta \cdot \Delta(\text{Yield spread})_{i,t-1} + \gamma \cdot \Delta(\widehat{AvgBidAsk})_{i,t+1} + \delta \cdot \Delta(\text{Controls})_{i,t} + \epsilon_{i,t}.$$

The control variables are the logarithm of the average trade size, the number of trades, and the logarithm of amount outstanding. Further, we employ rating dummies to control for credit risk. We calculate a bond's expected bid-ask spreads for month $t+1$ using either the naïve benchmark or the forecasting model. The naïve forecast for a bond's bid-ask spread in the next month equals the realization of today. The forecasting model is described in Section 2.3 and includes only variables that are in the restricted set of candidate predictors, i.e., directly related to liquidity (see Section 2.4). We measure bond liquidity using the average bid-ask spread measure. We winsorize changes in yield spreads and changes in expected bid-ask spreads at the 1% and 99% level. Differences in the out-of-sample mean square errors (MSE) are compared using the test statistic in Harvey, Leybourne, and Newbold (1997). Standard errors are clustered by bond and month and t-statistics are given in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Intercept	0.0034	0.0044	0.0091
	(0.12)	(0.16)	(0.34)
$\Delta(\text{Yield spread})_{i,t-1}$	0.1760*	0.1749*	0.1667*
	(1.86)	(1.86)	(1.84)
$\Delta(\widehat{AvgBidAsk})^{\text{naïve}}_{i,t+1}$		0.0636***	
		(3.69)	
$\Delta(\widehat{AvgBidAsk})^{\text{fc}}_{i,t+1}$			0.4534***
			(3.49)
$\Delta(\text{Trade size})_{i,t}$	-0.0138***	-0.0061	0.0020
	(-4.08)	(-1.31)	(0.29)
$\Delta(\text{Trades})_{i,t}$	0.0068***	0.0066***	0.0060***
	(3.71)	(3.67)	(3.52)
$\Delta(\text{Amount outstanding})_{i,t}$	0.1347	0.1358	0.1437
	(1.17)	(1.20)	(1.32)
$\Delta(\text{Rating dummies})_{i,t}$	Yes	Yes	Yes
R^2_{adj}	0.0624	0.0657	0.0808
$\Delta(R^2_{adj})$		0.0033	0.0184/0.0151
MSE	0.942	0.939	0.918
$\Delta(\text{MSE})$		-0.003***	-0.024***/-0.021***
		(5.57)	(14.29)/(17.56)
Observations		459,479	

Table 5: **Corporate bond fund flow regression**

This table reports results of the panel regressions of fund flows on expected liquidity changes of Section 3.1. We measure a fund's current and expected liquidity as the value-weighted average bid-ask spread across the bonds in the fund's portfolio. Expected liquidity is calculated using the restricted forecasting model of Section 2 and expected changes are relative to the current liquidity. Alpha is the intercept of a regression of excess fund returns on excess corporate bond and equity market returns. The dummy variable $\mathbb{1}_{\{\text{Alpha}<\theta\}}$ equals 1 if alpha is below a threshold θ, where we use 0, the upper quartile ($q_{75\%}$), the median ($q_{50\%}$), and the lower quartile ($q_{25\%}$) of all negative alphas. The quartiles correspond to monthly alphas of -5 bps, -12.5 bps, and -30 bps. We include lagged flow, the (natural) logarithm of total net assets, the logarithm of fund age in years, and net expense ratio as controls. We further employ an indicator variable that equals 1 if the fund charges rear load fees and 0 otherwise. The unit of observation is fund share class. We cluster standard errors by fund share class and include month fixed effects. t-statistics and, for cumulated effects, F-statistics are given in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

	$\theta = 0$		$\theta = q_{75\%}$		$\theta = q_{50\%}$		$\theta = q_{25\%}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ExpLiqChange		0.0003		0.0004		0.0007		0.0005
		(0.50)		(0.79)		(1.05)		(0.93)
ExpLiqChange × $\mathbb{1}_{\{\text{Alpha}<\theta\}}$		-0.0071***		-0.0083***		-0.0096***		-0.0151***
		(-3.96)		(-4.13)		(-4.28)		(-5.14)
FundLiq	0.0747	0.0289	0.1618	0.1158	0.2278*	0.1890	0.1085	0.0616
	(0.53)	(0.20)	(1.21)	(0.84)	(1.77)	(1.43)	(0.88)	(0.49)
FundLiq × $\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.6082***	-0.6557***	-0.7053***	-0.7356***	-1.0131***	-1.0848***	-1.2084***	-1.3802***
	(-4.09)	(-4.40)	(-4.63)	(-4.79)	(-6.06)	(-6.41)	(-5.89)	(-6.61)
$\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0088***	-0.0080***	-0.0070***	-0.0062***	-0.0035**	-0.0022	0.0001	0.0029
	(-6.41)	(-5.82)	(-4.82)	(-4.24)	(-2.08)	(-1.25)	(0.03)	(1.26)
Lagged flow	0.1514***	0.1513***	0.1517***	0.1516***	0.1519***	0.1518***	0.1525***	0.1523***
	(18.25)	(18.25)	(18.28)	(18.27)	(18.29)	(18.28)	(18.31)	(18.30)
TNA	0.0002	0.0003	0.0003*	0.0003*	0.0004**	0.0004**	0.0004**	0.0004**
	(1.46)	(1.49)	(1.89)	(1.91)	(2.21)	(2.19)	(2.12)	(2.07)
Age	-0.0207***	-0.0207***	-0.0209***	-0.0209***	-0.0211***	-0.0211***	-0.0210***	-0.0210***
	(-30.57)	(-30.60)	(-30.79)	(-30.80)	(-30.96)	(-30.95)	(-30.95)	(-30.94)
Expense	0.0009	0.0010	0.0013	0.0014	0.0009	0.0009	-0.0002	-0.0002
	(0.63)	(0.70)	(0.88)	(0.92)	(0.60)	(0.59)	(-0.12)	(-0.14)
Rear load	-0.0135***	-0.0136***	-0.0137***	-0.0138***	-0.0140***	-0.0140***	-0.0137***	-0.0137***
	(-8.16)	(-8.21)	(-8.30)	(-8.34)	(-8.42)	(-8.43)	(-8.26)	(-8.26)
R^2_{adj}	0.0639	0.0640	0.0637	0.0637	0.0635	0.0635	0.0630	0.0630
Cum. effect ExpLiqChange		-0.0069***		-0.0079***		-0.0090***		-0.0146***
		(15.66)		(16.43)		(17.04)		(25.35)
Cum. effect FundLiq	-0.5335***	-0.6267***	-0.5435***	-0.6198***	-0.7853***	-0.8958***	-1.0999***	-1.3186***
	(16.10)	(21.64)	(14.88)	(18.25)	(24.14)	(29.30)	(30.86)	(41.06)
Econ. sign. ExpLiqChange		-0.16%		-0.19%		-0.26%		-0.45%
Econ. sign. FundLiq	-0.24%	-0.28%	-0.25%	-0.28%	-0.36%	-0.41%	-0.54%	-0.65%
Observations	223,622	223,586	223,622	223,586	223,622	223,586	223,622	223,586

Table 6: **Corporate bond fund flow regression - Expected liquidity deterioration vs. improvement**

This table reports results of the panel regressions of fund flows on expected liquidity deterioration and expected liquidity improvement. We measure a fund's current and expected liquidity as the value-weighted average bid-ask spread across the bonds in the fund's portfolio. Expected liquidity is calculated using the restricted forecasting model of Section 2 and expected changes are relative to the current liquidity. To distinguish between the effect of expected liquidity deterioration and improvement, we use two dummy variables. $\mathbb{1}_{\{\text{ExpLiqChange}\leq 0\}}$ equals 1 if liquidity is expected to improve and $\mathbb{1}_{\{\text{ExpLiqChange}>0\}}$ is 1 if liquidity is expected to deteriorate. Alpha is the intercept of a regression of excess fund returns on excess corporate bond and equity market returns. The dummy variable $\mathbb{1}_{\{\text{Alpha}<\theta\}}$ equals 1 if alpha is below a threshold θ, where we use 0, the upper quartile ($q_{75\%}$), the median ($q_{50\%}$), and the lower quartile ($q_{25\%}$) of all negative alphas. The quartiles correspond to monthly alphas of -5 bps, -12.5 bps, and -30 bps. We include lagged flow, the (natural) logarithm of total net assets, the logarithm of fund age in years, and net expense ratio as controls. We further employ an indicator variable that equals 1 if the fund charges rear load fees and 0 otherwise. The unit of observation is fund share class. We cluster standard errors by fund share class and include month fixed effects. t-statistics are given in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

	$\theta = 0$	$\theta = q_{75\%}$	$\theta = q_{50\%}$	$\theta = q_{25\%}$
ExpLiqChange $\times\, \mathbb{1}_{\{\text{ExpLiqChange}\leq 0\}}$	-0.0031	-0.0040	-0.0036	-0.0019
	(-0.62)	(-0.80)	(-0.74)	(-0.58)
ExpLiqChange $\times\, \mathbb{1}_{\{\text{ExpLiqChange}\leq 0\}} \times \mathbb{1}_{\{\text{Alpha}<\theta\}}$	0.0019	0.0056	0.0049	-0.0070
	(0.34)	(1.08)	(0.98)	(-0.48)
ExpLiqChange $\times\, \mathbb{1}_{\{\text{ExpLiqChange}>0\}}$	0.0004	0.0007	0.0009	0.0007
	(0.80)	(1.07)	(1.25)	(1.16)
ExpLiqChange $\times\, \mathbb{1}_{\{\text{ExpLiqChange}>0\}} \times \mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0095***	-0.0127***	-0.0142***	-0.0157***
	(-4.86)	(-5.58)	(-5.65)	(-5.01)
FundLiq	0.0128	0.0923	0.1674	0.0610
	(0.09)	(0.66)	(1.25)	(0.48)
FundLiq $\times\, \mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.6550***	-0.7336***	-1.0856***	-1.3622***
	(-4.39)	(-4.78)	(-6.43)	(-6.30)
$\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0075***	-0.0055***	-0.0013	0.0030
	(-5.44)	(-3.69)	(-0.75)	(1.30)
Controls	Yes	Yes	Yes	Yes
R^2_{adj}	0.0640	0.0637	0.0636	0.0630
Observations	223,586	223,586	223,586	223,586

Table 7: **Robustness: Forecast performance**

This table reports root mean mean square errors (RMMSE) for the forecasting model and the naïve benchmark. In Panel A, we employ the random forest model of Section 4.1 as forecasting model and measure bond liquidity using the average bid-ask spread measure. The naïve forecast for a bond's liquidity in the next month equals the realization of today. In Panel B, we use the size-adapted average bid-ask spread of Section 4.2 to measure a bond's liquidity and calculate expected size-adapted bid-ask spreads employing the forecasting model of Section 2.3. The full model includes all candidate predictors from Section 2.2, whereas the restricted model only includes variables that are directly related to liquidity (see Section 2.4).

Panel A: Random forest model				
	Full model		Restricted model	
	RMMSE	Δ to naïve	RMMSE	Δ to naïve
Forecasting model	0.73	-20.22%	0.80	-19.30%
Naïve benchmark model	0.91		0.99	
Observations	230,790		511,465	
Panel B: Size-adapted liquidity measure				
	Full model		Restricted model	
	RMMSE	Δ to naïve	RMMSE	Δ to naïve
Forecasting model	0.79	-22.21%	0.76	-21.57%
Naïve benchmark model	1.02		0.97	
Observations	230,790		511,465	

Table 8: **Robustness: Yield spread regression**
This table reports results of the panel regressions of yield spread changes on changes in expected bid-ask spreads

$$\Delta(\text{Yield spread})_{i,t} = \alpha + \beta \cdot \Delta(\text{Yield spread})_{i,t-1} + \gamma \cdot \Delta(\widehat{AvgBidAsk})_{i,t+1} + \delta \cdot \Delta(\text{Controls})_{i,t} + \epsilon_{i,t}.$$

for the robustness checks of Sections 4.1 and 4.2. The control variables are the logarithm of the average trade size, the number of trades, and the logarithm of amount outstanding. Further, we employ rating dummies to control for credit risk. We calculate a bond's expected bid-ask spreads for month $t+1$ using either the naïve benchmark or the forecasting model. The naïve forecast for a bond's liquidity in the next month equals the realization of today. In Panel A, we employ the random forest model of Section 4.1 as forecasting model and measure bond liquidity using the average bid-ask spread measure. For convenience, we reprint specifications (1) and (2) from Table 4. In Panel B, we use the size-adapted average bid-ask spread of Section 4.2 to measure a bond's liquidity and calculate expected size-adapted bid-ask spreads employing the forecasting model of Section 2.3. Again, specification (1) is repeated from Table 4. Forecasting models include only variables that are in the restricted set of candidate predictors, i.e., directly related to liquidity (see Section 2.4). We winsorize changes in yield spreads and changes in expected (size-adapted) bid-ask spreads at the 1% and 99% level. Differences in the out-of-sample mean square errors (MSE) are compared using the test statistic in Harvey, Leybourne, and Newbold (1997). Standard errors are clustered by bond and month and t-statistics are given in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Panel A: Random forest model

	(1)	(2)	(3)
Intercept	0.0034	0.0044	0.0088
	(0.12)	(0.16)	(0.33)
$\Delta(\text{Yield spread})_{i,t-1}$	0.1760*	0.1749*	0.1670*
	(1.86)	(1.86)	(1.85)
$\Delta(\widehat{AvgBidAsk})^{\text{naïve}}_{i,t+1}$		0.0636***	
		(3.69)	
$\Delta(\widehat{AvgBidAsk})^{\text{fc}}_{i,t+1}$			0.4429***
			(3.44)
$\Delta(\text{Trade size})_{i,t}$	-0.0138***	-0.0061	0.0022
	(-4.08)	(-1.31)	(0.32)
$\Delta(\text{Trades})_{i,t}$	0.0068***	0.0066***	0.0059***
	(3.71)	(3.67)	(3.49)
$\Delta(\text{Amount outstanding})_{i,t}$	0.1347	0.1358	0.1737
	(1.17)	(1.20)	(1.56)
$\Delta(\text{Rating dummies})_{i,t}$	Yes	Yes	Yes
R^2_{adj}	0.0624	0.0657	0.0811
$\Delta(R^2_{adj})$		0.0033	0.0187/0.0154
MSE	0.942	0.939	0.915
$\Delta(\text{MSE})$		-0.003***	-0.027***/-0.024***
		(5.57)	(19.62)/(20.04)
Observations		459,479	

Table 8 continued

Panel B: Size-adapted liquidity measure			
	(1)	(2)	(3)
Intercept	0.0034	0.0051	0.0155
	(0.12)	(0.19)	(0.62)
Δ(Yield spread)$_{i,t-1}$	0.1760*	0.1745*	0.1599*
	(1.86)	(1.87)	(1.86)
$\Delta(\widehat{AvgBidAsk})^{\text{naïve}}_{i,t+1}$		0.1112***	
		(3.74)	
$\Delta(\widehat{AvgBidAsk})^{\text{fc}}_{i,t+1}$			0.9300***
			(3.79)
Δ(Trade size)$_{i,t}$	-0.0138***	-0.0178***	-0.0207***
	(-4.08)	(-5.91)	(-6.84)
Δ(Trades)$_{i,t}$	0.0068***	0.0066***	0.0060***
	(3.71)	(3.71)	(3.64)
Δ(Amount outstanding)$_{i,t}$	0.1347	0.1357	0.1458
	(1.17)	(1.21)	(1.40)
Δ(Rating dummies)$_{i,t}$	Yes	Yes	Yes
R^2_{adj}	0.0624	0.0700	0.1080
$\Delta(R^2_{adj})$		0.0076	0.0456/0.0380
MSE	0.942	0.933	0.882
Δ(MSE)		-0.009***	-0.060***/-0.051***
		(8.23)	(22.80)/(27.40)
Observations		459,479	

Table 9: **Robustness: Corporate bond fund flow regression**

This table reports results of the robustness checks for the impact of expected liquidity deterioration on corporate bond fund flows of Sections 4.1, 4.2, and 4.3. We measure a fund's current and expected liquidity as the value-weighted average across the bonds in the fund's portfolio. In Panel A, we employ the random forest model of Section 4.1 as forecasting model and measure bond liquidity using the average bid-ask spread measure. For convenience, specifications (1), (3), (5), and (7) are repeated from Table 5. In Panel B, we use the size-adapted average bid-ask spread of Section 4.2 to measure a bond's liquidity and calculate expected size-adapted bid-ask spreads employing the forecasting model of Section 2.3. Forecasting models include only variables that are in the restricted set of candidate predictors, i.e., directly related to liquidity (see Section 2.4). Expected liquidity changes are relative to the current liquidity. In Panel C, we employ the same specification as in Table 5, but include a fund's performance over the last year (alpha) and an interaction term as additional explanatory variables. Alpha is the intercept of a regression of excess fund returns on excess corporate bond and equity market returns. The dummy variable $\mathbb{1}_{\{\text{Alpha}<\theta\}}$ equals 1 if alpha is below a threshold θ, where we use 0, the upper quartile ($q_{75\%}$), the median ($q_{50\%}$), and the lower quartile ($q_{25\%}$) of all negative alphas. The quartiles correspond to monthly alphas of -5 bps, -12.5 bps, and -30 bps. We include lagged flow, the (natural) logarithm of total net assets, the logarithm of fund age in years, and net expense ratio as controls. We further employ an indicator variable that equals 1 if the fund charges rear load fees and 0 otherwise. The unit of observation is fund share class. We cluster standard errors by fund share class and include month fixed effects. t-statistics and, for cumulated effects, F-statistics are given in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level.

Panel A: Random forest model

	$\theta = 0$		$\theta = q_{75\%}$		$\theta = q_{50\%}$		$\theta = q_{25\%}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ExpLiqChange		-0.0001		0.0000		0.0001		0.0001
		(-0.20)		(0.12)		(0.31)		(0.19)
ExpLiqChange × $\mathbb{1}_{\{\text{Alpha}<\theta\}}$		-0.0043***		-0.0046***		-0.0047***		-0.0066***
		(-3.99)		(-4.13)		(-4.12)		(-5.22)
FundLiq	0.0747	0.0342	0.1618	0.1206	0.2278*	0.1942	0.1085	0.0723
	(0.53)	(0.24)	(1.21)	(0.89)	(1.77)	(1.48)	(0.88)	(0.58)
FundLiq × $\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.6082***	-0.6495***	-0.7053***	-0.7223***	-1.0131***	-1.0498***	-1.2084***	-1.2996***
	(-4.09)	(-4.36)	(-4.63)	(-4.71)	(-6.06)	(-6.23)	(-5.89)	(-6.27)
$\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0088***	-0.0081***	-0.0070***	-0.0064***	-0.0035**	-0.0027	0.0001	0.0017
	(-6.41)	(-5.88)	(-4.82)	(-4.39)	(-2.08)	(-1.57)	(0.03)	(0.76)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2_{adj}	0.0639	0.0639	0.0637	0.0637	0.0635	0.0635	0.0630	0.0630
Cum. effect ExpLiqChange		-0.0044***		-0.0045***		-0.0046***		-0.0066***
		(18.84)		(18.92)		(18.47)		(28.76)
Cum. effect FundLiq	-0.5335***	-0.6153***	-0.5435***	-0.6017***	-0.7853***	-0.8556***	-1.0999***	-1.2274***
	(16.10)	(21.35)	(14.88)	(17.64)	(24.14)	(27.71)	(30.86)	(37.09)
Econ. sign. ExpLiqChange		-0.15%		-0.17%		-0.20%		-0.35%
Econ. sign. FundLiq	-0.24%	-0.28%	-0.25%	-0.28%	-0.36%	-0.40%	-0.54%	-0.61%
Observations	223,622	223,586	223,622	223,586	223,622	223,586	223,622	223,586

Table 9 continued

Panel B: Size-adapted liquidity measure								
	$\theta = 0$		$\theta = q_{75\%}$		$\theta = q_{50\%}$		$\theta = q_{25\%}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ExpLiqChange		0.0016		0.0018*		0.0021*		0.0018*
		(1.56)		(1.72)		(1.87)		(1.76)
ExpLiqChange $\times \mathbb{1}_{\{\text{Alpha}<\theta\}}$		-0.0063***		-0.0064***		-0.0066***		-0.0064***
		(-4.58)		(-4.51)		(-4.53)		(-4.43)
FundLiq	0.0002	0.0001	0.0014	0.0013	0.0022*	0.0023*	0.0015	0.0016
	(0.12)	(0.07)	(0.98)	(0.96)	(1.66)	(1.69)	(1.13)	(1.20)
FundLiq $\times \mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0034**	-0.0039***	-0.0039***	-0.0043***	-0.0055***	-0.0061***	-0.0074***	-0.0083***
	(-2.41)	(-2.73)	(-2.64)	(-2.92)	(-3.19)	(-3.52)	(-3.39)	(-3.75)
$\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0110***	-0.0103***	-0.0098***	-0.0090***	-0.0079***	-0.0069***	-0.0047*	-0.0033
	(-8.20)	(-7.67)	(-6.85)	(-6.29)	(-4.46)	(-3.85)	(-1.92)	(-1.36)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2_{adj}	0.0638	0.0639	0.0636	0.0636	0.0634	0.0634	0.0628	0.0629
Cum. effect ExpLiqChange		-0.0048***		-0.0046***		-0.0045***		-0.0047***
		(23.59)		(21.95)		(20.99)		(20.03)
Cum. effect FundLiq	-0.0032**	-0.0038***	-0.0025*	-0.0030**	-0.0033*	-0.0039**	-0.0060***	-0.0067***
	(5.67)	(7.54)	(2.99)	(4.02)	(3.58)	(4.74)	(7.35)	(9.07)
Econ. sign. ExpLiqChange		-0.16%		-0.17%		-0.20%		-0.28%
Econ. sign. FundLiq	-0.16%	-0.18%	-0.12%	-0.15%	-0.16%	-0.19%	-0.31%	-0.34%
Observations	223,622	223,586	223,622	223,586	223,622	223,586	223,622	223,586

Table 9 continued

Panel C: Alpha as continuous variable

	$\theta = 0$		$\theta = q_{75\%}$		$\theta = q_{50\%}$		$\theta = q_{25\%}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ExpLiqChange		0.0002		0.0003		0.0005		0.0003
		(0.31)		(0.57)		(0.81)		(0.59)
ExpLiqChange $\times \mathbb{1}_{\{\text{Alpha}<\theta\}}$		-0.0055***		-0.0067***		-0.0083***		-0.0133***
		(-3.15)		(-3.48)		(-3.86)		(-4.65)
FundLiq	-0.0793	-0.1162	-0.0129	-0.0566	0.0422	0.0034	-0.0853	-0.1320
	(-0.55)	(-0.79)	(-0.09)	(-0.40)	(0.32)	(0.03)	(-0.69)	(-1.04)
FundLiq $\times \mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.2270	-0.2773*	-0.3350**	-0.3658**	-0.6854***	-0.7570***	-0.8667***	-1.0383***
	(-1.49)	(-1.81)	(-2.17)	(-2.35)	(-4.10)	(-4.48)	(-4.23)	(-4.98)
$\mathbb{1}_{\{\text{Alpha}<\theta\}}$	-0.0091***	-0.0085***	-0.0071***	-0.0065***	-0.0031*	-0.0020	0.0017	0.0042*
	(-6.67)	(-6.17)	(-4.91)	(-4.45)	(-1.85)	(-1.16)	(0.74)	(1.80)
Alpha	0.5771***	0.5767***	0.7133***	0.7147***	0.8432***	0.8395***	1.1012***	1.0939***
	(3.80)	(3.80)	(4.75)	(4.75)	(5.74)	(5.71)	(7.91)	(7.86)
$(\text{Alpha} - \theta) \times \mathbb{1}_{\{\text{Alpha}<\theta\}}$	0.4416*	0.3946*	0.1576	0.1125	-0.1883	-0.2331	-0.5482***	-0.6203***
	(1.93)	(1.75)	(0.73)	(0.53)	(-0.96)	(-1.22)	(-2.94)	(-3.44)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2_{adj}	0.0644	0.0644	0.0641	0.0641	0.0639	0.0639	0.0636	0.0637
Cum. effect ExpLiqChange		-0.0053***		-0.0064***		-0.0079***		-0.0130***
		(10.04)		(11.88)		(14.17)		(21.27)
Cum. effect FundLiq	-0.3063**	-0.3935***	-0.3479**	-0.4224***	-0.6432***	-0.7536***	-0.9520***	-1.1703***
	(5.14)	(8.13)	(5.98)	(8.29)	(16.31)	(20.89)	(23.14)	(32.34)
Econ. sign. ExpLiqChange		-0.12%		-0.16%		-0.23%		-0.40%
Econ. sign. FundLiq	-0.14%	-0.18%	-0.16%	-0.19%	-0.30%	-0.35%	-0.47%	-0.58%
Observations	223,622	223,586	223,622	223,586	223,622	223,586	223,622	223,586