ESSAYS ON FINANCIAL LIQUIDITY AND RISK

Andrew C. Robinson

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS

Adviser: Wei Xiong

June 2012
© Copyright by Andrew C. Robinson, 2012.

All rights reserved.
Abstract

This thesis is a collection of essays on financial liquidity and risk. The first two essays investigate the liquidity, liquidity premia, and liquidity risk premia of corporate bonds. The third essay examines the risk exposures of hedge fund strategies. The first two essays are single-authored, while the third is coauthored with William Kinlaw, John Papp, and David Turkington.

The first essay examines the liquidity and liquidity risk premia of investment-grade corporate bonds. I build on standard continuous-time structural credit models by incorporating an illiquid secondary market for bonds and by allowing this illiquidity to co-vary with Markov risk regimes. Then, using TRACE corporate bond transaction data from 2003 to 2011, two alternative measures of illiquidity, and consumption-liquidity regimes inferred from the data, I show that liquidity and liquidity risk have significant explanatory power in bond yield spreads, both in a reduced-form regression analysis, and in a structural model-calibration analysis. This effect is present both within and across bond rating classes, and is substantially stronger in the period following the 2007-2009 financial crisis.

The second essay explores whether on-the-run corporate bonds are more liquid than comparable off-the-run bonds, and whether this liquidity carries a premium. Using corporate bond issuance and transaction data, I measure the monthly liquidities of 504 bonds from 36 distinct issuers. In a panel regression analysis, I show that on-the-run bonds are substantially more liquid than comparable off-the-run bonds, where liquidity is measured by volume traded, frequency of transactions, and implied bid-ask spread. To investigate the existence of a liquidity premium, I compare yields of portfolios containing on-the-run bonds with the yields of portfolios containing off-the-run bonds. I find weak evidence that an average spread of 5-8 basis points exists. More interestingly, I find that the spread explodes in October 2008, potentially reflecting a flight-to-liquidity phenomenon.
that is consistent with other time-series studies of aggregate liquidity.

The third essay examines non-linearities in the risk exposures of hedge fund returns. We analyze the exposures of six broadly defined hedge-fund strategies to a diverse collection of risk factors, including equity markets, commodities, Fama-French factors, credit risk, and options. The analysis extends the existing literature in two ways: by looking at monthly hedge fund strategy returns through the 2008 crisis period, and by analyzing recently available daily hedge fund strategy returns. In contrast to previous research, we find that for monthly returns, at-the-money and close to at-the-money put options on the S&P500 index are not correlated with hedge funds strategies. However, many of the hedge fund monthly returns, as well as many of the daily returns, are negatively correlated with deeper out-of-the-money puts.
Acknowledgements

I would like to thank my adviser Wei Xiong for his consistent guidance and encouragement throughout my graduate studies. Many other people provided invaluable advice, support, and instruction. They include Sylvain Chassang, Dong Beom Choi, Hyun Soo Choi, Thomas Eisenbach, Ricardo Gambirasio, Mike Geruso, Harrison Hong, Jakub Jurek, Zhuan Pei, Martin Schmalz, Hyun Song Shin, and Adam Zawadowski. John Papp deserves a special mention for engaging me in a countless number of discussions and debates. Participants in the Princeton Finance Workshop contributed many helpful suggestions. I gratefully acknowledge financial support from the Princeton University Graduate School and the Princeton Department of Economics. Finally, I am deeply indebted to my family, especially my mother Juliann, my father Jamie, and my sister Lana, for their love and moral support throughout my time at Princeton.
## Contents

**Abstract** iii  
**Acknowledgements** v  
**Table of Contents** vi  
**List of Tables** ix  
**List of Figures** x  

1 Modeling and Measuring Corporate Bond Liquidity Risk Premia 1  

1.1 Introduction 1  
1.2 Related Literature 4  
1.3 The Model 5  

1.3.1 Economy and Firm $P$ and $Q$ Dynamics 6  
1.3.2 Firm Capital Structure 9  
1.3.3 Bond Illiquidity 10  
1.3.4 Equilibrium Debt Valuation 12  
1.3.5 Equilibrium Equity Valuation and Spreads 14  
1.4 Calibration 17  

1.4.1 Data Description and Summary 18  
1.4.2 Economy, Firm, and Bond Parameters 20  
1.4.3 Liquidity Measures 22  
1.4.4 Regime Inference 25  
1.4.5 Regression Analysis 27  
1.4.6 Model Performance 29  
1.5 Conclusion 32
1.A Model Solution Details ................................................. 33
  1.A.1 Debt Problem ....................................................... 33
  1.A.2 Equity Problem ................................................... 36
1.B Kalman Filter Details .................................................. 44
  1.B.1 Notation ............................................................... 44
  1.B.2 Estimation ............................................................ 46

2 On-the-run and Off-the-run Corporate Bond Liquidity 57
  2.1 Introduction ............................................................ 57
  2.2 Data Description ...................................................... 58
    2.2.1 Bonds, Issuers, and Transactions ............................... 59
    2.2.2 Measuring Liquidity ............................................. 61
  2.3 Determinants of Bond Liquidity ................................... 62
    2.3.1 Bond Age and Liquidity ......................................... 62
    2.3.2 Time Series Trends in Aggregate Liquidity ................... 65
    2.3.3 Are On-the-run Bonds More Liquid? A Panel Analysis ............ 66
  2.4 Analyzing the Spread between On-the-run and Off-the-run Corporate Bond Yield Spreads ............................................. 68
    2.4.1 Constructing a Spread Trade ................................... 69
    2.4.2 The On-the-run Off-the-run Corporate Bond Spread .............. 71
  2.5 Conclusion ............................................................. 74
  2.A The Roll Illiquidity Measure ....................................... 75

3 Non-linear Hedge Fund Risk Exposures 94
  3.1 Introduction ........................................................... 94
  3.2 Monthly Data Sources ................................................. 96
    3.2.1 Hedge Fund Indices ............................................. 96
<table>
<thead>
<tr>
<th>Table Number</th>
<th>Table Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bond Summary Statistics</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>Bond Yield Spreads and Liquidity</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>Yield Spreads and Crisis Liquidity</td>
<td>51</td>
</tr>
<tr>
<td>4</td>
<td>Bond, Issuer, and Transaction Summary Statistics</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>List of Bond Issuers</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>Monthly Bond Liquidity Measures</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>Liquidity Panel Regressions, No Fixed Effects</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>Liquidity Panel Regressions, Issuer Fixed Effects</td>
<td>81</td>
</tr>
<tr>
<td>9</td>
<td>Liquidity Panel Regressions, with Bond and Month Fixed Effects</td>
<td>82</td>
</tr>
<tr>
<td>10</td>
<td>The On-the-run Off-the-run Corporate Bond Spread Trade</td>
<td>83</td>
</tr>
<tr>
<td>11</td>
<td>Hedge Fund Index and Risk Factor Monthly Summary Statistics</td>
<td>111</td>
</tr>
<tr>
<td>12</td>
<td>Risk Factor Analysis, Partial Sample: Monthly Returns, January 1997 to June 2000</td>
<td>112</td>
</tr>
<tr>
<td>13</td>
<td>Risk Factor Analysis Full Sample: Monthly Returns, January 1997 to March 2011</td>
<td>113</td>
</tr>
<tr>
<td>15</td>
<td>Hedge Fund Index and Risk Factor Daily Summary Statistics</td>
<td>115</td>
</tr>
<tr>
<td>16</td>
<td>Risk Factor Analysis with Out-of-the-Money Put: Daily Returns, April 1 2003 to April 29 2011</td>
<td>116</td>
</tr>
<tr>
<td>17</td>
<td>Risk Factor Analysis with Out-of-the-Money Put and Selected Risk Factors: Daily Returns</td>
<td>117</td>
</tr>
</tbody>
</table>
List of Figures

1. Macroeconomic Variables and the Crisis .......................... 52
2. Monthly Stress Regime Probabilities Implied by the Kalman Filter .... 53
3. Calibrated Model Performance, Full Sample (1781 Bonds) ............. 54
4. Calibrated Model Performance, A-rated bonds (784 bonds) ........... 55
5. Calibrated Model Performance, Pre- and Post- Crisis ................ 56
6. Monthly Par Volume Traded by Age .................................. 84
7. Monthly Transactions by Age ........................................... 85
8. Monthly Turnover by Age ................................................. 86
9. Monthly Roll Illiquidity by Age ....................................... 87
10. Aggregate Monthly Volume per Bond by Month ....................... 88
11. Aggregate Monthly Transactions per Bond by Month .................. 89
12. Aggregate Monthly Roll Illiquidity by Month ......................... 90
13. Average Yield Spreads, by Month ..................................... 91
14. Spreads on the On-the-run / Off-the-run Long-Short Trades by Month 92
15. Lehman On-the-run and Off-the-run Price and Yield Spreads, August-September 2008 ........................................................... 93
16. Cumulative Monthly Hedge Fund Index Returns ....................... 118
17. Monthly HFRI Returns vs S&P500 Returns ............................ 119
1 Modeling and Measuring Corporate Bond Liquidity Risk Premia

1.1 Introduction

A substantial literature is devoted to explaining corporate yield spreads using structural credit models. Expected losses to default can typically only explain a small fraction of spreads; for example, Huang and Huang (2003) show that while average Baa-spreads are about 150 bps historically, a standard structural credit model, calibrated to match historical default rates and other parameters, predicts that such a bond should only require a spread of 32 bps. At least two additional factors have significant explanatory power: the price of consumption, and illiquidity. The consumption regime literature\(^1\) has shown that default risk can potentially explain spreads if the model accounts for the fact that default is more likely to occur in an expensive state of the world. Likewise, the illiquidity literature\(^2\) has shown that expected liquidity costs, under various measures, can also explain spreads.

This paper addresses the intersection of these two avenues of research by considering an environment where both fundamentals and liquidity are subject to regimes. The relevance of such an environment is readily apparent: asset illiquidity, including corporate bond illiquidity, was a defining feature of the recent crisis. Intuition suggests that a bond that is illiquid at the very time when its default risk is highest and most costly to investors should command an additional premium, during both “normal” regimes and “stress” regimes. My research agenda is thus to address the following questions: 1) Do bonds that become more illiquid in a stress regime trade at higher spreads in the normal regime, controlling for average illiquidity? 2) Can a structural credit model that incorporates regime-varying illiquidity capture this variation? 3) Can the model explain spread variation within bond rating classes as well as variation across rating classes?

\(^1\)For example, Chen et al. (2009), Bhamra et al. (2009), and Chen (2011)

\(^2\)For example, Renault and Ericsson (2002), Chen et al. (2007), Bao et al. (2010)
My analysis contains both theoretical and empirical components. I begin by building on standard continuous-time structural credit models (for example, Leland (1994)) in which default, which occurs when firm assets fall to an endogenously determined threshold, is the only source of risk and the only determinant of spreads. By incorporating an illiquid bond market and a probability that bondholders will be forced to liquidate (as in He and Xiong (2011), for example), I add expected liquidity costs as an additional component of spreads. By including an aggregate consumption process whose drift and volatility follow Markov regimes (as in Chen (2011)), I account for the fact that default is more likely to occur when it is more costly to bondholders (which is realized in spreads as a default premium). Finally, by allowing asset liquidity to also depend on the regimes, I allow liquidity to affect spreads through two additional risk channels: 1) the direct cost associated with paying a higher liquidity cost in a regime in which consumption is more valuable, and 2) the indirect cost stemming from higher “rollover” costs faced by equity.

I then employ TRACE data from April 2003 to March 2011, which contains detailed transaction-level information on corporate bond yields. Two additional variables are essential to my analysis: bond liquidity, and consumption-liquidity regimes. I construct two distinct measures of individual-bond liquidity: the Amihud (2002) price-impact measure, which assumes that a bond is more illiquid if its price moves more in response to a given volume traded; and the Roll (1984) bid-ask bounce measure, which assumes that a bond is more illiquid if has a greater transitory component. I then infer liquidity-consumption regimes from the monthly aggregate Amihud and Roll measure time-series and S&P500 return series, using a Markov-switching assumption and the Markov state-space Kalman filter approach due to Hamilton (1989). With liquidity measures and consumption-liquidity regimes in hand, I regress spreads of investment-grade bonds on liquidity and regime-specific liquidity and a variety of controls, including rating, age, time-to-maturity, and average volume traded. I show that illiquid bonds do indeed trade at higher spreads: for
example, increasing a bond’s Amihud liquidity by one standard deviation increases its spread by 8 bps, while increasing its Roll measure by one standard deviation increases its spread by 30 bps. Furthermore, bonds that were more illiquid during the stress regime have higher average spreads, controlling for average illiquidity: for example, increasing a bond’s stress-regime Amihud liquidity by one standard deviation increases its average spread by 39 bps, while increasing its stress-regime Roll measure by one standard deviation increases its average spreads by 18 bps.

To evaluate the ability of the model to capture this effect of liquidity regimes on spreads, I perform a bond-by-bond calibration of the model, using economy, firm, and bond-specific parameters. I show that the interaction between liquidity and consumption regimes does indeed have additional explanatory power, above and beyond the separate effects of the price of consumption and average illiquidity costs. Furthermore, a significant fraction of this result persists when the bonds are compared within rating classes. The model thus helps explain some of the observed within-rating spread variation that is correlated with illiquidity risk, something that a standard structural credit model that is calibrated to match a single representative firm cannot do. Finally, I show that this ability of the model to capture spread variation has improved dramatically in the post-crisis period. This suggests that bond liquidity, and more specifically, the covariance of liquidity and the price of consumption, may be more relevant for bond pricing in the wake of a crisis in which asset liquidity was a defining characteristic.

The rest of the paper is organized as follows. In Section 1.2, I review the relevant literature, including papers on structural credit models, spread decomposition, regimes, asset liquidity, and liquidity risk. In Section 1.3, I construct my model, solve it, and derive a set of testable predictions. In Section 1.4, I describe my data sets, construct liquidity measures and regimes to calibrate the model, present regression results, and evaluate the model performance. In Section 1.5 I conclude.
1.2 Related Literature

Corporate bond pricing is the subject of an extensive literature, both theoretical and empirical. The theoretical literature has traditionally focused on default risk as the primary determinant of yield spreads, both from a “structural” perspective (for example, Merton (1974), Black and Cox (1976), Leland (1994), Longstaff and Schwartz (1995), Leland and Toft (1996)) as well as from a “reduced-form” perspective (for example, Duffee (1999), Duffie and Singleton (1999)). However, a number of empirical papers have argued that observed spreads are far too large to account for simple expected default losses: for example, Huang and Huang (2003) show that while average Baa-spreads are about 150 bps historically, a standard structural credit model, calibrated to match historical default rates and other parameters, predicts that such a bond should only require a spread of 32 bps.

A recent strand of literature has shown that structural default models can in fact explain spreads, if the model accounts for the fact that default is more likely in states of the world in which it is also more costly. For example, Chen et al. (2009) show that, in a continuous-time model with a time-varying pricing kernel and an exogenous default boundary, historical default rates can be consistent with historically observed spreads, if either the pricing kernel negatively co-varys with firm assets, or the pricing kernel positively co-varys with the default boundary. Similarly, in contemporaneous works, Chen (2011) and Bhamra et al. (2009) develop general equilibrium models with endogenous capital structures, endogenous default boundaries, and Markov consumption regimes, and show that the risk premium generated from regime switches can explain spreads.

An alternative explanation for the discrepancy between spreads and expected default losses is illiquidity. A substantial literature examines the pricing of liquidity and liquidity risk, for both corporate bonds and other assets such as equities. In any early work, Amihud and Mendelson (1986) show that stocks that are more illiquid, as proxied for
by larger bid-ask spreads, earn higher returns. Amihud (2002) shows that the expected illiquidity of equities, measured by price-impact, is priced in both time-series and cross-sectional dimensions. Pastor and Stambaugh (2003) and Acharya and Pedersen (2005) establish a canonical liquidity-CAPM framework for equity returns, and demonstrate that stocks whose returns are correlated with aggregate liquidity measures receive higher returns. Using this framework, Lin et al. (2011) and de Jong and Driessen (2005) show that liquidity risk is also priced in corporate bond returns. Acharya et al. (2010) take the analysis a step further by showing that the pricing of corporate bond liquidity betas is regime-switching.

The two papers most similar to mine belong to Chen et al. (2007) and Bao et al. (2010). These papers directly examine the effect of individual-level bond liquidity on spreads. Chen et al. (2007) employ bond prices from 1995 to 2003 and a variety of liquidity measures, including the Pastor and Stambaugh (2003) measure and a zero-returns measure, and show that more illiquid bonds trade at higher spreads. Bao et al. (2010) use TRACE data through 2009 and the bid-ask bounce illiquidity measure due to Roll (1984) and similarly show that more illiquid bonds command higher spreads. The first part of my analysis effectively replicates these results: I use TRACE data through 2011 and show that average illiquidity, as measured by the Roll (1984) as well as the Amihud (2002) measure, explains spreads. The second part of my analysis is the marginal contribution: I show that bonds that were more illiquid during the crisis command an additional premium, controlling for average illiquidity, and that this premium persists post-crisis.

1.3 The Model

I construct a structural credit model with the following goals in mind. First, the model should be consistent with existing structural credit models, to facilitate comparisons of
model predictions. Second, the model should have a concrete and intuitive representation of bond liquidity. Third, the model should incorporate economic regimes in a tractable manner. Fourth, the model should endogenously deliver bond default rates, prices, and spreads as a function of liquidity, regime, and other variables. And finally, the model should be analytically solvable to the greatest extent possible. I begin with a standard structural credit model (see, for example, Merton (1974), Black and Cox (1976), Leland (1994), Leland and Toft (1996)), in which firm assets follow an exogenous stochastic process, and default endogenously occurs when the value of equity drops to zero. I extend this by incorporating an illiquid secondary bond market (see He and Xiong (2011)); in effect, bond illiquidity is modeled as a chance that a bondholder might be shocked into costly liquidation. Finally, because we are interested in the effect of liquidity in different regimes, I place the above firm in an economic environment in which both the price of consumption and the cost of liquidation are subject to Markov-switching regimes (see Chen (2011)).

This section is organized as follows. In Section 1.3.1, I present the economic environment of the model: the consumption-liquidity regimes, the $P$-measure dynamics of the economy and the representative firm, and the representative-agent utility function that links the $P$ and $Q$ measures. In Section 1.3.2, I present the capital structure of the firm and the optimization problems of debt and equity holders. In Section 1.3.3, I incorporate bond illiquidity as a transaction cost. Finally, in Sections 1.3.4 and 1.3.5, I solve for the debt and equity holder value functions and thus derive bond spreads.

1.3.1 Economy and Firm $P$ and $Q$ Dynamics

The model consists of a representative firm embedded in a continuous-time economy that is subject to Markov regime switches. The economic regime $S$ controls both the drift $\mu_Y^P(S)$ and the volatility $\sigma_Y^P(S)$ of the consumption process, which is assumed to follow
a Brownian motion under the physical measure, conditional on the regime:

\[
\frac{dY_t}{Y_t} = \mu_Y(S)dt + \sigma_Y(S)dZ_t^Y.
\]

The subscript \(Y\) signifies that the drift and volatility refer to the economy-wide consumption process \(Y\), while the superscript \(P\) signifies that these are moments of the \(P\)-measure distribution. The economic regime also controls asset liquidity; I discuss this in detail later in Section 1.3.3. Though this type of model can be adapted to include many regimes (for example, both Bansal and Yaron (2004) and Chen (2011) use nine regimes), I restrict my model to two, a “normal” regime and a “stress” regime, which are sufficient to analyze the effect of correlated rises in the price of consumption and cost of illiquidity. In the normal regime \(S = 1\), consumption growth has a higher expectation and lower volatility. In the stress regime \(S = 2\), consumption growth has a lower expectation and a higher volatility. The economy switches from the normal regime to the stress regime with Poisson intensity \(\eta_{12}^P\) and from the stress regime to the normal regime with intensity \(\eta_{21}^P\) (again, under the \(P\)-measure).

I focus the analysis on a single firm with an exogenous asset portfolio. Firm asset returns are assumed to be exogenously driven by both the systematic and an idiosyncratic Brownian motion:

\[
\frac{dX_t}{X_t} = \mu_X(S)dt + \sigma_X(S)dZ_t^X + \beta_X\sigma_Y(S)dZ_t^Y,
\]

where \(\mu_X(S)\) is the expected return of firm assets in regime \(S\), \(\sigma_X(S)\) is the idiosyncratic asset volatility in regime \(S\), and \(\beta_X\sigma_Y(S)\) is the systematic asset volatility in regime \(S\).

The above return processes are given under the \(P\)-measure. \(P\)-measure distributions are useful because they can easily be calibrated to match moments of observed historical data. Of course, to derive the values of debt and equity we need to account for investor
risk-aversion by using the $Q$-measure. The literature broadly handles this in two ways. The partial equilibrium approach is simply to assume a risk-neutral measure and start from there, ignoring any relation to historically observed physical distributions. In this case, specifying the aggregate consumption process is unnecessary, as the exogenously specified $Q$-measure firm asset process is sufficient to price the debt and equity securities. The second, general equilibrium approach, is to assume a representative agent utility function and use it to derive a $Q$-measure from a historically calibrated $P$-measure. In this case, the $P$-distribution of the aggregate consumption process are critical to deriving the state-price densities used to value debt and equity. For example, Chen et al. (2009) use a Campbell-Cochrane pricing kernel to derive $Q$ from $P$. Similarly, Bansal and Yaron (2004) and Chen (2011) employ Epstein and Zin (1989) preferences in an economy with Markov switching regimes. I follow the second approach. Under this assumption, firm assets follow a geometric Brownian motion with drift $\mu^Q_X(S)$ and volatility $\sigma^Q_X(S)$:

\[
\frac{dX_t}{X_t} = \mu^Q_X(S)dt + \sigma^Q_X(S)dZ_t,
\]

where the $Q$-measure drift accounts for the systemic risk of firm assets. Similarly, the Markov switching intensities $\eta^Q_{12}$ and $\eta^Q_{21}$ under $Q$ account for the fact that consumption is more expensive in the stress regime than in the normal regime. The specific equations relating $P$ and $Q$ measure moments are given in the appendix.

Before moving to the firm’s capital structure, it is worth discussing the limitations of the model so far. Three observations warrant discussion. First, firm assets are assumed to follow a completely exogenous stochastic process. Firms do not adjust either the size or composition of their assets in response to changes in their balance sheet or the regime. While this assumption is standard, it precludes analysis of complex interactions between a firm’s balance sheet, its debt value, and the economy. Second, the assumption of only
two regimes, which permits relatively convenient closed-form solutions for debt and equity values, prevents the calibrated model from generating $P$-distribution moments that accurately match historical data. Doing this requires more regimes; for example, Chen (2011) uses nine regimes. The final observation concerns the parametric assumptions of the regime-switching model itself. A regime-switching model, while more flexible than a single-regime model, can still have trouble fitting observed $P$ and $Q$ moments. More flexible parametric models, for example, (continuous) stochastic volatility models, or jump-diffusion models, have better results. The most flexible approach makes no parametric assumptions. For example, Breeden and Litzenberger (1989) develop a method for deriving the entire state price distribution from the prices of traded equity index options.

1.3.2 Firm Capital Structure

The focus of the model is on the secondary-market valuation of firm debt. To facilitate analysis, I make the following assumptions about the firm's capital structure. Like He and Xiong (2011) and Leland (1994), I assume that the firm issues only two types of securities, debt and equity, and that the amount of debt and its characteristics are exogenous and constant: at any given time, the firm is assumed to have a continuum of bonds outstanding with aggregate principal $p$ and aggregate coupon flow $c$. To further simplify the analysis, I assume that bonds have a random maturity structure: they mature independently with Poisson intensity $\phi$. This is similar to the approaches used by Leland (1994) and Chen (2011), who consider perpetual bonds. Removing time-dependency allows us to solve for the value of debt and equity in closed-form as a function of the default boundary. The random maturity intensity parameter $\phi$ implies an average maturity of $\frac{1}{\phi}$, which I later calibrate to match actual maturities. This stationary maturity approach is in contrast to He and Xiong (2011), who analyze a bond with time-varying maturity and use a Laplace transformation to solve for debt and equity values in closed-
The additional complexity of Markov regimes in my model makes a stationary debt structure a welcome simplification.

The assumption of a continuum of independently maturing bonds implies that over any small time interval $dt$, the total principal of bonds maturing equals $\phi p dt$. To maintain its debt structure, the firm must roll over these bonds by issuing new ones (with identical maturity intensities and coupon rates) at market value. Let $D$ denote the market value of debt with principal $p$, coupon $c$, and intensity $\phi$. The difference between the principal payment $\phi p dt$ and the revenue generated by issuing new debt $\phi D dt$ is the rollover gain or loss. Like He and Xiong (2011), I assume that any gain is immediately paid out to equity holders, while any loss is covered by issuing new equity (and diluting existing equity). Rollover losses therefore create additional downside risk: a negative shock to firm assets will reduce debt values and increase rollover losses, further decreasing equity values.

As is standard in structural valuation models, equity holders endogenously default when firm assets fall to an endogenously-determined threshold $X(S)$. Clearly this threshold will depend on the state $S$. Intuitively, the threshold in the low consumption regime $X(2)$ will be higher than the threshold in the high consumption regime $X(1)$; we will indeed verify that this is the case. In the event of default, equity holders walk away with nothing, while debt holders share a fraction of the remaining assets $lX(S)$, where the fraction $1 - l$ is lost to bankruptcy costs. The analysis is thus centered on the valuation of debt, the valuation of equity, and the default thresholds.

1.3.3 Bond Illiquidity

While there are many distinct empirical measures of asset liquidity, the general consensus for representing liquidity in an asset pricing model is to model it as a fractional transaction cost (see, for example, Acharya and Pedersen (2005), He and Xiong (2011)). I follow
existing literature by modeling bond illiquidity as a cost incurred when debt holders are shocked into liquidating their holdings. More specifically, I assume that with Poisson intensity $\xi$, debt holders receive a liquidity shock and must liquidate all of their bonds at a fraction $1 - k(S)$ of their current market value. By allowing the liquidation cost $k$ to potentially depend on the state $S$ we can analyze the effect of procyclical liquidity on defaults and spreads.

Bond liquidity affects debt values (and spreads) through two distinct channels. First there is the direct channel: a bond that is more illiquid will be more costly to liquidate if its holder is forced to sell, and thus should trade at a discount. Additionally, a bond that is more illiquid in bad states of the world should expect to trade at an additional liquidity risk discount, because liquidation is more costly in this regime. The second, indirect channel is known as the rollover risk channel: a bond that is more illiquid (or that is more illiquid in the low consumption regime) will command a lower price at issuance. Since equity holders must pay any difference between the face value of debt and its market value when rolling over maturing bonds, illiquidity will lower equity values and thus raise default thresholds, which of course feeds back into higher spreads.

In the following sections, I solve the model by jointly deriving debt values, equity values, and default thresholds. I show that bond illiquidity has the effects predicted above. Fixing other characteristics, a bond that has a higher $k$ will have a higher spread; moreover, both the direct trading-cost channel and the indirect rollover-risk channel significantly contribute. Second, fixing $k$, a bond that is more illiquid in the stress regime relative to the normal regime (ie, has a higher $k(2)$ and a lower $k(1)$) commands a significantly higher spread; and again, both channels contribute significantly.
1.3.4 Equilibrium Debt Valuation

The solution to this model is comprised of three elements: a debt value function, an equity value function, and a pair of default thresholds. I proceed as follows. First, I take the default thresholds as given and characterize and solve the debt holder’s valuation problem. Second, given the debt holder’s value function, I characterize and solve the equity holders’ valuation problem. The default threshold is thus the fixed point solution to this pair of problems. Finally, I derive bond spread.

Let $D(X, S)$ denote debt value as a function of firm assets $X$ and the consumption-liquidity regime $S$. Such a contingent claim in a continuous-time model is typically time-dependent and thus characterized by a partial differential equation (for example, the Black-Scholes model). However, because of the random-maturity structure in our model, debt values do not explicitly depend on $t$ itself; thus, given the Brownian dynamics of $X$ and the Markov dynamics of $S$, debt dynamics solve a pair of ordinary differential equations. For the high consumption regime $S_1$, $D(X, S_1)$ is defined on $X \geq \overline{X}_1$ and satisfies:

$$r_1 D(X, S_1) = c + \phi(p - D(X, S_1)) - \xi k D(X, S_1) +$$

$$+ \mu^Q_X(S_1)XD_X(X, S_1) + \frac{\sigma^Q_X(S_1)^2}{2}X^2 D_{XX}(X, S_1) +$$

$$+ \mathbb{1}_{X > \overline{X}_2} \eta^Q_{12}(D(X, S_2) - D(X, S_1)) +$$

$$+ \mathbb{1}_{X_2 > X > \overline{X}_1} \eta^Q_{12}(lX - D(X, S_1)).$$ (1)

The left-hand side is the debt value function multiplied by $r_1$, the risk-free rate in the high consumption state $S_1$. This is the required rate of return in $S_1$, under the risk-neutral measure. The right-hand side contains three sets of terms. The first set contains expected payouts today: $c$ is the coupon flow payment; $(p - D(X, S_1))$ is the principal payment, net of the value of the bond, which occurs when the bond matures with intensity
ϕ; and $k_1D(X, S_1)$ is the liquidity loss, which occurs with intensity $\xi$. The second set of terms, $\mu^Q_1(S_1)XD_1(X, S_1)$ and $\frac{\sigma^2_1(S_1)}{2}X^2D_{XX}(X, S_1)$, accounts for the risk-neutral drift and volatility of the Brownian asset dynamics. Finally, the third set of terms accounts for the Markov dynamics of the regime $S$. A regime switch occurs with intensity $\eta_{12}$, in which case, one of two things can happen. If $X > \overline{X}_2$, the firm remains above the state-2 default threshold, and does not default; in this case, debt value jumps down to $D(X, S_2)$. On the other hand, if $\overline{X}_2 > X > \overline{X}_1$, the regime switch puts the firm under its state-2 default threshold. The firm immediately defaults and debtholders recover a fraction $l$ of assets.

Similarly, debt value in the low-consumption state $S_2$ is defined on $X \geq \overline{X}_2$ and satisfies:

$$r_2D(X, S_2) = c + \phi(p - D(X, S_2)) - \xi k_2 D(X, S_2) +$$

$$+ \mu^Q_1(S_2)XD_1(X, S_2) + \frac{\sigma^2_1(S_1)}{2}X^2D_{XX}(X, S_2) +$$

$$+ \eta_{21}(D(X, S_1) - D(X, S_2)).$$

Nearly every term is an analogue to a term in (1), with one exception. Since $X \geq \overline{X}_2$ implies that $X \geq \overline{X}_1$, a regime switch from low to high will never cause an immediate default; instead, it will always cause debt values to jump to $D(X, S_1)$.

Four boundary conditions are required to solve this system of ODEs. The upper boundary conditions are standard debt upper boundary conditions: as firm assets approach infinity, debt values should remain bounded. The lower boundary conditions simply state that debt recovers a fraction $l$ of the assets in the event of default, in either regime:

$$\lim_{X \to \overline{X}_i} D(X, S_i) = l\overline{X}_i,$$
for \( i = 1, 2 \).

To solve this system I take a guess-and-verify approach, the details of which are given in Appendix 1.A. Given the default thresholds \( \overline{X}_i, i = 1, 2 \), debt values satisfy:

\[
D(X, S_1) = \begin{cases} 
    d_{11}X^{\alpha_1} + d_{13}X^{\alpha_3} + d_{15} & \text{if } X > \overline{X}_2 \\
    d_{21}X^{\beta_1} + d_{22}X^{\beta_2} + d_{25} + d_{26}X & \text{if } \overline{X}_2 > X > \overline{X}_1
\end{cases}
\]

\[
D(X, S_2) = \begin{cases} 
    d_{31}X^{\alpha_1} + d_{33}X^{\alpha_3} + d_{35} & \text{if } X > \overline{X}_2 \\
    lX & \text{if } \overline{X}_2 > X > \overline{X}_1
\end{cases}
\]

The expressions for coefficients \( d_i \) and exponents \( \alpha_i \) are also given in Appendix 1.A.

### 1.3.5 Equilibrium Equity Valuation and Spreads

I now compute equity values. Equity valuation is substantially more complex than debt valuation, because of the dependence of equity values on debt values through the rollover channel. I solve for equity values taking the default thresholds \( \overline{X}_i \) and the debt value functions \( D(X, S_i) \) as given. Like debt values, equity values will be the solution to a pair of ODEs, one for each regime. Let \( E(X, S) \) denote the value of equity as a function of firm assets and the state. For the high consumption regime \( S_1 \), \( E(X, S_1) \) is defined on \( X \geq \overline{X}_1 \) and satisfies:

\[
\begin{align*}
    r_1E(X, S_1) &= \delta X - (1 - \pi)c + \phi(D(X, S_1) - p) + \\
    &+ \mu_X^Q(S_1)XE_X(X, S_1) + \frac{\sigma_X^Q(S_1)^2}{2}XXE_{XX}(X, S_1) + \\
    &+ 1_{X > \overline{X}_2} \eta_{12}^Q(E(X, S_2) - E(X, S_1)) + \\
    &+ 1_{\overline{X}_2 > X > \overline{X}_1} \eta_{12}^Q(-E(X, S_1)).
\end{align*}
\]  

(2)
Like the debt value ODEs, the equity value ODEs can be decomposed into three sets of terms. The first set \( \delta X - (1 - \pi)c + \phi(D(X, S_1) - p) \) accounts for current expected payouts: \( \delta X - (1 - \pi)c \) is the dividend flow payout to equity holders, less the tax-adjusted coupon flow paid to bond holders; and \( \phi(D(X, S_1) - p) \) is the rollover gain or loss flow associated with rolling over principal \( \phi p \) at market value \( \phi D(X, S_1) \). The second set of terms accounts for the risk-neutral drift and volatility of the firm’s assets. And finally, the third set of terms accounts for expected losses that occur during a regime change, if the firm assets are above \( 1_{X > \overline{X}_2} \eta_{12}(E(X, S_2) - E(X, S_1)) \) or below \( 1_{X_2 > X > \overline{X}_1} \eta_{12}(-E(X, S_1)) \) the state-2 default threshold.

Equity value in the low consumption state is similarly defined on \( X \geq \overline{X}_2 \) and satisfies

\[
\begin{align*}
  r_2 E(X, S_2) &= \delta X - (1 - \pi)c + \phi(D(X, S_2) - p) + \\
  &+ \mu^Q_X(S_2) X E_X(X, S_2) + \frac{\sigma^Q_X(S_2)^2}{2} X^2 E_{XX}(X, S_2) + \\
  &+ \eta^Q_{21}(E(X, S_1) - E(X, S_2)).
\end{align*}
\]

Like debt values, equity value in the low consumption state is analogous to its value in the high consumption state, with the exception that a regime switch never triggers an immediate default.

These ODEs are accompanied by six boundary conditions. First, the upper boundary conditions simply state that, as assets approach infinity, \( E(X, S_1) \) and \( E(X, S_2) \) should be proportional to \( X \). This logically follows from the boundedness of debt values. The lower boundary conditions simply state that equity is worth zero at the default thresholds:

\[
\lim_{X \to \overline{X}_i} E(X, S_i) = 0,
\]

15
for $i = 1, 2$. Finally, two “smooth pasting” boundary conditions are required to pin down the optimal default thresholds (for example, see Dixit (2002)):

\[
\lim_{X \to X_i} E_X(X, S_i) = 0.
\]

Because of the dependence of equity values on debt values in the equity value ODEs, solving for the equity value functions is substantially more complicated than solving for debt. However, the time-homogeneity of the environment again allows a closed-form solution. Given the default thresholds, the equity value functions are given by:

\[
E(X, S_1) = \begin{cases} 
  e_{11}X^{\alpha_1} + e_{13}X^{\alpha_3} + e_{15} + e_{16}X + e_{17}X^{\alpha_7} + e_{19}X^{\alpha_9} & \text{if } X > \bar{X}_2; \\
  e_{21}X^{\beta_1} + e_{22}X^{\beta_2} + e_{25} + e_{26}X + e_{27}X^{\beta_7} + e_{28}X^{\beta_8} & \text{if } \bar{X}_2 > X > \bar{X}_1; 
\end{cases}
\]

\[
E(X, S_2) = \begin{cases} 
  e_{31}X^{\alpha_1} + e_{33}X^{\alpha_3} + e_{35} + e_{36}X + e_{37}X^{\alpha_7} + e_{39}X^{\alpha_9} & \text{if } X > \bar{X}_2; \\
  0 & \text{if } \bar{X}_2 > X > \bar{X}_1. 
\end{cases}
\]

The expressions for the coefficients $e_i$ and exponents $\beta_i$ and the details of the solution are given in Appendix 1.A. Finally, the default thresholds are found by applying the smooth pasting conditions to the equity value functions. Because of the functional form of the equity functions, it is not possible to analytically solve for the thresholds; however, it is straightforward to calculate them numerically as the solution to a fixed point problem. The details are also given in Appendix 1.A.

Our ultimate interest is in bond spreads, the difference between the bond yields and the yields on risk-free Treasuries. In this model, a bond’s yield is defined as its required rate of return, conditional on holding it until maturity (ie, no liquidation or default). Recall that the bond matures with intensity $\phi$, and that the regime switches from $i$ to $j$. 
with intensity $\eta_{ij}$. Given firm assets $X$ and regime $S_i$, yield $y$ satisfies

$$y(X, S_i)D(X, S_i) = c + \phi(p - D(X, S_i)) + \eta_{i,-i}(D(X, S_{-i}) - D(X, S_i)),$$

which has the solution

$$y(X, S_i) = \frac{c}{D(X, S_i)} + \phi\left(\frac{P}{D(X, S_i)} - 1\right) + \eta_{i,-i}\left(\frac{D(X, S_{-i})}{D(X, S_i)} - 1\right).$$

The yield spread is given by the difference between yield and the risk free rate, $y(X, S_i) - r_i$. In our analysis, we will be interested in the consequences of liquidity and regimes on normal-regime spreads $y(X, S_1)$, on stress-regime spreads $y(X, S_2)$, and on unconditional spreads

$$\bar{y}(X) = \frac{\eta_{21}}{\eta_{12} + \eta_{21}}y(X, S_1) + \frac{\eta_{12}}{\eta_{12} + \eta_{21}}y(X, S_2).$$

### 1.4 Calibration

Our goal is to investigate the spread implications of bond liquidity in a variety of circumstances. While most papers use a single representative firm to calibrate a structural credit model (using average firm characteristics), I employ a number of firm- and bond-specific parameters to perform a bond-by-bond calibration of the model. This allows me to better assess the quality of the model; for example, while a standard calibration is evaluated by its ability to match averages, my model can also be evaluated by its ability to match cross-sectional variation, such as within-rating spread variation. This section is organized as follows. In Section 1.4.1, I describe the data sets that I employ. In Section 1.4.2, I discuss how the data is used to generate many of the economy, firm, and bond parameters. In Section 1.4.3, I characterize two measures of bond liquidity and present their empirical properties. In Section 1.4.4, I describe my method of inferring normal and stress regimes using historical liquidity and consumption data. In Section 1.4.5, I
perform a series of regressions of spreads on liquidity and regime-liquidity to show that bonds that become more illiquid in the stress regime trade at higher spreads, controlling for average liquidity. Finally, in Section 1.4.6, I present the calibrated model results.

1.4.1 Data Description and Summary

I employ several data sets to calibrate the model. The main data set I use is FINRA’s TRACE, which contains bond transaction data beginning in 2002. To obtain bond characteristics I merge this data set with the Mergent Fixed Income Securities database, which contains issue details and ratings. I merge these data sets with CRSP to obtain equity dynamics. To generate yield spreads I use U.S. Treasury rates published by the Federal Reserve Board. Finally, to calibrate leverage parameters I match the bond data to firm balance sheet data using COMPUSTAT.

TRACE is the bond transaction database for the Financial Industry Regulatory Authority, a non-governmental regulator of the securities industry. TRACE contains transaction-level data for over-the-counter corporate bond trades. TRACE reporting began in July 2002, covering only large issues, but has covered 99% of all bond transactions since February 2005. TRACE data only contains transaction-relevant variables such as clean price and par-value traded. For bond characteristics, I merge the TRACE transaction data with the Mergent Fixed Income Securities. The Mergent FISD Issue database contains bond issue characteristics, such as volume, maturity, coupon size and frequency, and callability, for over 140,000 bonds. Additionally, the Mergent FISD Ratings database contains initial and subsequent Moody’s, S&P, and Fitch ratings for the vast majority of these bonds. I match bonds from TRACE and FISD using the 9-character CUSIPs that uniquely identify North American securities. Finally, to fully calibrate the model, I need to employ two additional data sets. CRSP contains daily stocks prices and is used to generate firm asset volatility and systematic volatility. COMPUSTAT contains quar-
terly balance sheet data and is used to generate firm leverage.³

I begin with the entire set of bond transactions recorded in TRACE, from its inception in July 2002 until March 2011. Since TRACE reporting was restricted to a much smaller selection of bonds prior to April 2003, I drop all transactions before this date. After removing duplicate transactions and transactions that were clearly misreported, I collapse the data to daily volumes and volume-weighted prices⁴. I then merge the TRACE data with the FISD issue and ratings data and eliminate all putable bonds, sinking-fund bonds, perpetual bonds, and bonds with floating rates. Because callable bonds comprise nearly 50% of the sample, I do not eliminate these bonds, but instead include a dummy for callability in my regressions. I also eliminate all bonds I cannot match with issue or rating data. I restrict the sample to bonds that enter the data set as investment-grade (which account for over 75% of over-the-counter transactions anyway). Finally, I eliminate bonds that do not satisfy criteria needed to construct meaningful liquidity measures; these criteria are described in the following section. For the most part, this means eliminating bonds that are traded extremely rarely or infrequently.

Table 1 summarizes my sample, which contains 1781 bonds over the 96-month period between April 2003 and March 2011. There are 68,370 bond-month observations, indicating that the average bond in our sample traded in just under 40% of the months included in our time-frame. On average, a bond is traded on 19 days in months in which it is traded at least once. Bonds rated A comprise 53% of the sample, followed by Aa rated bonds (23%), Baa rated bonds (20%), and Aaa rated bonds (4%). Callable bonds comprise 48% of the sample. Financial bonds comprise 23% of the sample. The average bond in the sample is about 3.25 years old and has about 5.81 years left until maturity.

³Only about 60% of bond issuers can be reliably matched to COMPUSTAT; rather than drop 40% of my observations, I assume that the average leverage of this sample is unbiased.

⁴Though the analysis focuses on monthly spreads, daily prices are needed to construct monthly Amihud and Roll liquidity measures.
The average offering amount is about $1 billion, and the average monthly volume traded is $84 million. Bonds rated Baa have a much higher turnover rate (25%) than bonds rated A (7%) or Aa (7%). The average spread of all bonds is 165 basis points. The average spreads of Aaa and Aa, A, and Baa bonds are 121, 145, and 251 bps respectively. Of course, there is dramatic variation in spreads over the time frame we consider; average spreads are as low as 59 bps (August 2005), and as high as 557 bps (October 2008). Panel (c) of Figure 1 shows the variation of average spreads over time.

1.4.2 Economy, Firm, and Bond Parameters

In this section I describe how I calibrate the majority of the model parameters to match my data. These parameters are broadly divided into three categories: economy (physical drifts $\mu^p(S)$ and volatilities $\sigma^p(S)$ of the consumption process; risk-neutral drifts $\mu^q(S)$; and physical and risk-neutral regime-switching intensities $\eta^p_{ij}$ and $\eta^q_{ij}$), firm (physical and risk-neutral drifts of firm assets $\mu^p_X(S)$ and $\mu^q_X(S)$; physical idiosyncratic asset volatilities $\sigma^p_X(S)$; and asset beta $\beta_X$), and bond (principal $p$; coupon $c$; maturity intensity $\phi$; bankruptcy recovery rate $1 - l$; corporate tax rate $\tau$; liquidity shock frequency $\xi$; and liquidation costs $k(S)$).

First, consider the physical and risk-neutral dynamics of the consumption process. As described in the previous section, the challenge here is to relate risk-neutral dynamics, which are used to derive spreads and are not directly observable, to historical physical dynamics, which can be observed. The standard approach is to choose physical parameters that generate moments consistent with the data, and then assume a representative agent utility function that generates risk-neutral dynamics from the physical dynamics. However, standard time-separable preferences are not capable of generating reasonable risk-neutral dynamics from historical physical dynamics (see Mehra and Prescott (1985), for example). One solution, employed by Bansal and Yaron (2004) and Chen (2011) and...
others, is to use Epstein and Zin (1989) preferences, which, combined with Markov consumption regimes, can simultaneously generate reasonable physical and risk-neutral dynamics. I follow that approach here. The calibration task is thus to match the parameters of the physical consumption process, \( \mu_Y(S), \sigma_Y(S), \eta_{12}, \) and \( \eta_{21} \), for \( S = 1, 2 \), with the moments of the observed aggregate consumption process, as proxied by the S&P500. To do so requires us to identify the normal and stress regimes; I describe how this is done in Section 1.4.4. These parameters are then combined with standard Epstein and Zin (1989) preferences, as in Bansal and Yaron (2004) and Chen (2011), to generate reasonable Q-measure dynamics.\(^5\)

Next, consider the \( P \)- and \( Q \)-measure firm asset dynamics. Asset dynamics are only directly observable with quarterly frequency (from COMPUSTAT, for example), which is far from ideal for a reasonably accurate estimate. However, equity dynamics are obviously observable with daily (or higher) frequency, and, assuming that firms maintain relatively constant leverage, and that firms are far enough away from their default boundaries that equity dynamics are linear in asset dynamics, we can approximate the drifts and total volatilities of firm assets by the respective equity parameters, divided by leverage. Likewise, asset beta is given by equity beta, divided by leverage, which allows us to separately identify idiosyncratic and systematic volatilities. The risk-neutral firm dynamics are then derived in the same manner as the risk-neutral economy dynamics.

Finally, consider the remaining bond-specific parameters. The coupon rate \( c/p \) is matched to the bond-specific coupon rate given in Mergent’s FISD. Maturity intensity \( \phi \), which implies and expected maturity of \( 1/\phi \) years, is matched to bond tenor, as observed in TRACE. The bankruptcy recovery rate \( 1 - l \) is set equal to 50 percent, which is the

\(^5\)At least, as reasonable as one can generate with only two regimes; as Bansal and Yaron (2004) observe, an ideal calibration contains at least nine regimes. Unfortunately, incorporating nine regimes in my model would dramatically increase the complexity of the analytical model solution and obscure much of the intuition behind results.
value that is widely used in the literature as the recovery rate of investment-grade bond defaults. The corporate tax rate is set equal to 35 percent. The liquidity shock frequency $\xi$, which implies an average annual turnover rate of $1/\xi$, is matched to observed annual bond turnover, under the assumption that all transactions are motivated by a liquidity shock to exactly one of the trading parties. Finally, the liquidation cost $k(S)$ is set to match implied average illiquidity costs; this is described in detail in the following section.

### 1.4.3 Liquidity Measures

Asset liquidity is consistent with at least several definitions. Generally, an asset is said to be more illiquid if it has greater transaction costs (implicit or explicit), or if trading a given quantity results in greater price movements (either transitory, or persistent). The literature utilizes a number of measures of asset liquidity. Perhaps the simplest is the quoted bid-ask spread (for example, see Amihud and Mendelson (1986)). Roll (1984) introduces a bid-ask “bounce” measure, which effectively is a bid-ask spread implied by transitory price movements; this measure is used by Bao et al. (2010) in their analysis of corporate bond illiquidity. Lesmond et al. (1999) construct a zero-return measure, based on the premise that a more illiquid bond will trade on fewer days and thus have more zero-return days. The measure introduced by Amihud (2002) is based on the premise that a more illiquid bond will have a higher absolute return, relative to volume traded (persistent price impact). The Pastor and Stambaugh (2003) measure is predicated on the idea that a more illiquid bond will have greater price reversals following large trades (transitory price impact).

I employ two distinct measures of liquidity in my analysis.\textsuperscript{6} The first, developed by Amihud (2002), is the average ratio of absolute daily returns to daily volume. I hence-

\textsuperscript{6}I will consistently refer to both measures as “liquidity” measures, rather than “illiquidity” measures, even though both measures are increasing in \textit{illiquidity}.
forth refer to this as the Amihud liquidity measure $Amihud^i$ of bond $i$. Suppose that bond $i$ is traded on $D_i$ days; let $R_d^i$ and $V_d^i$ denote the daily return and volume on day $d$. Then $Amihud^i$ is given by:

$$Amihud^i = \frac{1}{D_i} \sum_{d=1}^{D_i} \frac{|R_d^i|}{V_d^i}.$$ 

Bonds that have a higher absolute return, relative to volume traded, have a higher Amihud liquidity measure, and are said to be more illiquid. The Amihud measure is used by Acharya and Pedersen (2005), Acharya et al. (2010), and Lin et al. (2011) in their construction of liquidity betas.

The second measure I employ was developed by Roll (1984) and used by Bao et al. (2010), among others. It is based on the premise that a bond will trade at a higher price if the transaction is buyer-initiated, and a lower price if the transaction is seller-initiated (i.e., there is a bid-ask spread). Under the assumption that the transaction-initiator is independent of the initiator of the previous transaction, we would expect price changes to have negative serial covariance due to this bid-ask bounce. This is in contrast to a random walk model, where price changes are independent. A greater covariance (in absolute terms) implies a larger bid-ask bounce and hence a more illiquid asset. The Roll measure $Roll^i$ of bond $i$ is defined by

$$Roll^i = -\text{Cov}(\Delta P_d^i, \Delta P_{d+1}^i),$$

where $P_d^i$ is the price of bond $i$ on day $d$, and $\Delta P_d^i = P_d^i - P_{d-1}^i$. Bonds whose prices “bounce” more have a higher Roll liquidity measure, and are said to be more illiquid.

Following the literature, I eliminate bonds according to several criteria to improve the quality of our liquidity measures. The Roll bid-ask bounce measure requires bond prices on consecutive days. I follow Bao et al. (2010) and eliminate bonds that trade for
less than one year, bonds that do not have at least 10 consecutive-trade-day pairs, and bonds that trade on less than 75% of the business days between their first trade date and their last trade date. Finally, I follow Lin et al. (2011) and Winsorize both the Amihud and Roll liquidity measures using the 1st and 99th percentiles of the distribution each month to reduce the impact of outliers.

Before summarizing the liquidity measures, it is necessary to discuss their interpretations. Recall that we have modeled illiquidity as a fractional transaction cost. Our two liquidity measures, on the other hand, are price impact and bid-ask bounce measures, with no immediate transaction-cost interpretations. Luckily, under the assumptions of the Roll model, the Roll measure implies a proportional transaction cost of $2\sqrt{\text{Roll}_i}$.

Inferring transaction costs from the Amihud liquidity measure is more complicated, because it is not derived from a trading model. To remedy this, Acharya and Pedersen (2005) perform an affine transformation on their Amihud liquidity measure so that its first and second moments in their data match the respective moments in a different transaction-cost data set (in their case, the bid-ask spreads in Chalmers and Kadlec (1998)). I replicate this procedure, but instead use the Roll measure in our own data set. Thus, the mean and standard deviation of the Amihud liquidity measure in our full sample are artificially set equal to the mean and standard deviation of the Roll liquidity measure (see table 1). Nevertheless, the correlation between the two measures is only 0.60, implying that each measure partially captures a different aspect of illiquidity.

The Amihud and Roll liquidity measures are summarized in the last columns of table 1. The Roll liquidity measure has a mean of 0.41 and a standard deviation of 0.43 in the entire sample, corresponding to an average transaction cost of about 131 bps. By construction, the Amihud liquidity measure has the same mean and standard deviation as the Roll measure, over the entire sample. Higher rated bonds are significantly more liquid under both measures: the average transaction cost for Aa rated bonds is about 104
bps, compared to 151 bps for Baa rated bonds. Liquidity also varies dramatically over
time, as shown in Panels (a) and (b) of figure 1, which plot monthly aggregate liquidity
(measured by unweighted averages of Amihud and Roll measures). Both aggregate Ami-
hud and Roll measures peak during October of 2008, at the height of the panic; Amihud illiquidity during this month is 3.38, almost 7 times its average, and Roll illiquidity during this month is 2.43, almost 6 times its average.

1.4.4 Regime Inference

The defining feature of my model is the presence of two consumption-liquidity regimes.
The next step in the analysis is thus identifying the normal regime and the stress regime
in the data; clearly the stress regime should coincide with the 2007-2009 crisis that spans
the middle of our data set. Identifying the regimes allows me to calibrate the model to
match regime-specific parameters.

In my model, a consumption-liquidity stress regime is characterized by two things:
higher average illiquidity, and a lower drift and more volatile consumption process. With
this in mind, I use three time series to estimate the regimes: the aggregate monthly Ami-
hud and Roll liquidity measures, constructed by taking unweighted averages each month,
and the S&P500 return series. Figure 1 plots aggregate monthly Amihud and Roll liquid-
ity and S&P500 returns between April 2003 and March 2011. Casual inspection suggests
that the period between mid-late 2007 and mid-2009 is a good candidate for a stress-
regime: in this period, illiquidity is dramatically higher and more volatile under both
measures, and S&P500 returns are also lower and more volatile.

To make the regimes more precise, I employ a Markov state-space Kalman filter (see,
for example, Hamilton (1989) and Kim (1994)). Three assumptions form the basis of this
approach. First, I assume that there are two consumption-liquidity regimes, which follow
a Markov switching process. Second, I assume that both liquidity and consumption re-
turns follow state-dependent stochastic processes, whose functional forms but not parameters are exogenous. And third, I assume that the three observed time series are noisy estimates of the true liquidity and consumption processes. Under these assumptions, the model specification is summarized by three equations:

\[ S_t = N'S_{t-1}; \]  

\[ L_t = A(S_t)L_{t-1} + (I - A(S_t))D(S_t) + G(S_t)\nu_t; \]  

\[ Y_t = F(S_t)Illicit + \epsilon_t; \]

where \( N = \{\eta\}_{ij} \) is the Markov transition matrix for the state; \( L_t = (Illicit_t, Fundamental_t)' \) are the true, unobserved liquidity and consumption values at \( t \), \( Y_t = (Amihud_t, Roll_t, S&P500_t)' \) is the vector of signals at \( t \); \( A(S_t), D(S_t), G(S_t) \) and \( F(S_t) \) are parameter scalars and matrices; and \( \nu_t \) and \( \epsilon_t \) are error terms with mean zero and variance covariance matrices \( R \) and \( Q \) respectively. Equation 3 simply states that states transition according to the Markov matrix \( N \). Equation 4 states that the underlying processes follow AR(1) specifications, conditional on the state, with coefficients \( A(S_t) \) and long-run means \( D(S_t) \). Finally, equation 5 states that the aggregate Amihud and Roll measures and the returns of the S&P500 are noisy estimates of the true underlying liquidity and fundamental processes.

Estimation follows four steps. First, the Kalman filter is applied, giving us an estimate \( \hat{L}^{ij}_t \) of the underlying process \( L_t \) at each month \( t \), conditional on \( S_{t-1} = i, S_t = j \), and information up to \( t \). Second, these state-conditional estimates are approximated by unconditional estimates \( \hat{L}_t \), using the conditional state densities. Third, the parameters \( A(S_t), D(S_t), G(S_t), F(S_t), N, R, \) and \( Q \) are estimated via maximum likelihood. And finally, the Kalman smoother is applied, giving us period-by-period estimates of the un-
derlying liquidity and consumption processes and the probability of a stress regime using past, present, and future information. Details on the filter, the approximation procedure, and the parameter estimation are given in Appendix 1.B.

Figure 2 plots the monthly probability of a stress regime. As shown, this probability jumps sharply from 0.17 to 0.99 in August 2007, and it remains high until May 2009, when it falls from 0.90 to 0.15. We thus choose the 21-month period between August 2007 and April 2009 as the stress regime, and the remaining months in the sample as the normal regime. The parameters of the consumption dynamics are roughly matched to the average S&P500 drifts and volatilities during these periods, and the firm specific parameters are matched to equity dynamics during these periods, as described in Section 1.4.2. Finally, the liquidation costs $k(S)$ are matched to the firm-specific average values of the Amihud and Roll liquidity measures in these periods, as described in Section 1.4.3.

### 1.4.5 Regression Analysis

The model presumably predicts that bonds that become more illiquid in stress regimes trade should trade at higher spreads in both regimes. Before calibrating the model, I investigate whether this liquidity risk premium actually exists, using two sets of regressions. I first show that bond liquidity, as measured using the entire sample, has economically-significant power in spreads. To this end I perform monthly Fama-Macbeth regressions of spreads on Amihud and Roll liquidity measures and a standard set of controls, including age, tenor, log issuance size, log monthly volume, bond beta, a callability dummy, a financial dummy, and rating dummies. The Amihud and Roll liquidity measures here are constructed using the entire sample; the liquidity regimes are not used in this stage. I show that, as expected (eg, Chen et al. (2007), Bao et al. (2010)), bonds that are more illiquid also command higher spreads.

The second stage of the analysis directly addresses the main prediction of the model:
bonds that are more illiquid during a liquidity crisis should command higher spreads out of the crisis, controlling for average liquidity and other factors. For this purpose I construct “crisis” Amihud and Roll liquidity measures, using just observations between August 2007 and April 2009. I then perform monthly Fama-Macbeth regressions on controls, full-sample liquidity measures, and the crisis liquidity measures. I show that crisis-liquidity does indeed have statistically and economically significant explanatory power, above and beyond full-sample liquidity.

Table 2 presents the results of the first set of regressions. Using the entire sample of bond-month observations, I perform monthly Fama-Macbeth regressions of spreads on Amihud liquidity, Roll liquidity, age, tenor, log issuance size, log monthly volume, bond beta, a callability dummy, a financial dummy, and rating dummies. I use Newey and West (1987) standard errors to control for possible auto-correlation. Specifications (5) and (6) demonstrate that the Amihud and Roll liquidity measures do indeed have significant effects on spreads, even in the presence of other controls: Amihud liquidity has a statistically-significant coefficient of 18.98, while Roll liquidity has a statistically-significant coefficient of 56.39. To put these numbers in perspective, both the full-sample Amihud and the full-sample Roll measures have standard deviations of 0.43, implying that increasing a bond’s Amihud (Roll) liquidity by one standard deviation is associated with a spread increase of 8 (24) bps, holding other characteristics constant. To compare, Bao et al. (2010) use a similar set of regressors and TRACE data through 2009 and find that a one-standard-deviation increase in their Roll liquidity measure is associated with a spread increase of about 35 bps.

Table 3 presents the results of a second set of regressions. Columns (1) and (2) are replicated from Table 2 for comparison: they show the effect of the full-sample liquidity measures on spreads. Specifications (3) and (4) add an additional factor: “crisis” liquidities. These are bond-specific Amihud and Roll liquidity measures constructed using only
data between August 2007 and April 2009. The full-sample liquidity measures are also included in the regressions to control for average liquidity; recall that the main prediction of our model is that bonds that are more illiquid in poor states of the world should command higher spreads in all states, controlling for average liquidity. Table 3 shows that crisis liquidity does indeed command a significant spread premium, even when controlling for total liquidity. When included in separate regressions, the Amihud and Roll crisis liquidity measures have significant coefficients of 45.21 and 23.40. Taking into account that the two liquidity measures are more volatile during the crisis (standard deviations of 0.64 and 0.82 respectively), these results imply that increasing a bond’s crisis Amihud or Roll liquidity by one standard deviation increases its spread by 39 or 18 bps.

1.4.6 Model Performance

I now describe the performance of the calibrated model. A standard calibration approach compares observed spreads to the model-predicted spread of a single-representative bond whose parameters have been calibrated to match averages in the data. In contrast, my calibration approach delivers one predicted spread per bond. My analysis is thus concerned with not only how well the model delivers accurate average spreads, but also how well it captures the cross-sectional variation in spreads, both within and across ratings. I address the following questions. How well does the model perform in matching both average spreads, and the cross-sectional variation in spreads? How do the two alternative liquidity measures compare? How much of the model performance is due to the interaction between liquidity and consumption regimes? Does the model explain within-rating spread variation, or just across rating? How does the model performance vary across time?

I first consider the model’s performance in explaining the (time-series) average spreads of the full sample of 1781 bonds. Figure 3 plots the 1781 pairs of spreads and predicted
spreads; here, each bond’s spread is the time-series average. The (cross-sectional) average spread is 165 bps, with a standard deviation of 102 bps (recall that the sample only includes investment-grade bonds). Panel (a) plots the observed spreads versus the spreads predicted by the model, calibrated with the Amihud liquidity measure. The model performs reasonably well, predicting an average spread of 107 bps. A linear fit has a positive and significant slope of 0.23 and an $R^2$ of 0.19, implying that the model has some success in explaining cross-sectional spread variation. Similarly, Panel (b) plots the observed spreads versus the spreads predicted by the model calibrated with the Roll liquidity measure. The Roll measure does markedly better, delivering an average predicted spread of 127 bps, a positive and significant slope of 0.28, and an $R^2$ of 0.26.

To address the question of how much of this result is due to the interaction between liquidity and the consumption regimes, I calibrate a variant of the model in which bond-specific liquidity costs $k$ of each bond are constant across regimes. Panels (c) and (d) of Figure 3 plot actual spreads against the spreads predicted by this variant of the model, calibrated with Amihud and Roll measures respectively. This variant of the model does substantially worse in predicting both the average level of spreads, and their variation: the average predicted spreads are 72 and 93 bps respectively, with slopes of 0.13 and 0.18 and $R^2$s of 0.10 and 0.15. This implies that the interaction between liquidity and the price of consumption, which is the modeling innovation of this paper, is quantitatively significant.

Intuition suggests that spread differences across rating groups is responsible for at least part of this result. For example, the difference between the average spread of Aa-rated and the average spread of Baa-rated bonds is over 100 bps. The question is whether the model can explain within-rating variation, or whether it is limited to explaining across-rating variation. To investigate this question, I replicate the analysis using just A-rated bonds, which comprise over 40% of the sample (784 bonds), and have an average spread
of 145 bps. The answer is that the model does indeed capture some within-rating spread variation. Figure 4 plots the performance of the model in capturing spread variation within this rating class. Panel (a) plots actual spreads against the spreads predicted by the model calibrated with the Amihud measure. The model predicts an average spread of 95 bps, compared to the actual average spread of 145 bps. The slope of the linear regression line is 0.15, which is substantially smaller than the slope of the full-sample regression line, but still quite significantly different from zero. The $R^2$ is 0.14 (compared to 0.19 for the full sample), implying that at least part of the model’s success in explaining spread variation is within-rating. Similarly, Panel (b) plots the actual spreads versus spreads predicted by the model calibrated with Roll measure data. The average predicted spread is 115 bps, with a slope of 0.20 and an $R^2$ of 0.23. Finally, Panels (c) and (d) plot the performance of the model when liquidity is restricted to be constant across regimes (but allowed to varying across bonds). Again, the model performs substantially worse under this restriction, suggesting that the liquidity-consumption regime interaction is economically significant in explaining both across-rating and within-rating variation.

The final question I address is how the significance of the interaction between illiquidity and the price of consumption in bond pricing changes over time. To do this, I consider two separate periods, the pre-crisis period, defined as April 2003 to July 2007, and the post-crisis period, defined as May 2009 to March 2011. I calibrate the model twice, using bond data from these periods. I restrict the analysis to include only A-rated bonds and the Roll liquidity measure. Panels (a) and (b) of Figure 5 plot the performance of the model using pre-crisis data and the constant liquidity and regime-varying liquidity assumptions respectively. Panels (c) and (d) plot the performance of the model using post-crisis data and the constant liquidity and regime-varying liquidity assumptions. The graphs clearly show that the model performs dramatically better in the post-crisis period. The average pre-crisis spread for A-rated bonds was 76 bps, with a standard devi-
ation of only 35 bps. Both the constant-liquidity model variant and the regime-varying liquidity model variant predict average spreads of 45 bps, with \( R^2 \) of 0.04. In contrast, the average post-crisis spread was 129 bps, with a standard deviation of 70 bps, and the model predicts average spreads of 96 and 115 bps with constant and varying liquidity respectively. The \( R^2 \)s of 0.32 and 0.44 are substantially higher. These results suggest that bond liquidity, and more specifically, the covariance of liquidity and the price of consumption, may be more relevant in the wake of a crisis in which asset liquidity was widely viewed as one of the main culprits.

1.5 Conclusion

I conduct a theoretical and empirical investigation into the effects of liquidity and consumption-liquidity regimes on corporate bond yield spreads. The contributions of the paper are threefold. First, I provide an analytically tractable model, consistent with standard continuous-time structural credit models, that incorporates “stress” regimes characterized by a higher cost of consumption and higher asset illiquidity. Second, using bond transaction data from 2003-2011, I show that bonds that became more illiquid during the 2007-2009 crisis trade at higher spreads, holding average liquidity constant. And third, I perform a bond-by-bond calibration that allows me to evaluate the ability of the model to capture not only average spread levels, but spread variation, across rating classes, within rating classes, and over time. My analysis suggests that both corporate bond illiquidity and illiquidity during stress regimes have become increasingly relevant in bond pricing.
1.A Model Solution Details

1.A.1 Debt Problem

To solve the debt problem, I use a guess-and-verify approach. To simplify notation, let \(\mu_S\) denote \(\mu_X(S)\), \(\sigma_S\) denote \(\sigma_X(S)\), \(\eta_{ij}\) denote \(\eta_{ij}^Q\), and \(\bar{X}_S\) denote \(\bar{X}(S)\), for \(S, i, j = 1, 2\). Taking thresholds \(\bar{X}_1\) and \(\bar{X}_2\) as given for the moment, conjecture that the debt value functions take the forms:

\[
D(X, S_1) = \begin{cases} 
  d_{11}X^{\alpha_1} + d_{13}X^{\alpha_3} + d_{15} & \text{if } X > \bar{X}_2; \\
  d_{21}X^{\beta_1} + d_{22}X^{\beta_2} + d_{25} + d_{26}X & \text{if } \bar{X}_2 > X > \bar{X}_1;
\end{cases}
\]

\[
D(X, S_2) = \begin{cases} 
  d_{31}X^{\alpha_1} + d_{33}X^{\alpha_3} + d_{35} & \text{if } X > \bar{X}_2; \\
  lX & \text{if } \bar{X}_2 > X > \bar{X}_1.
\end{cases}
\]

To simplify notation, write

\[
\rho_1 = r + \phi + \xi k_1 + \eta_{12};
\]

\[
\rho_2 = r + \phi + \xi k_2 + \eta_{21};
\]

\[
A_0 = c + \phi p.
\]

The PDEs then imply the following. For \(X > \bar{X}_2\), in the high consumption state:

\[
\rho_1(d_{11}X^{\alpha_1} + d_{13}X^{\alpha_3} + d_{15}) = A_0 + \mu_1(d_{11}\alpha_1 X^{\alpha_1} + d_{13}\alpha_3 X^{\alpha_3}) + \\
+ \frac{\sigma_1^2}{2}(d_{11}\alpha_1(\alpha_1 - 1)X^{\alpha_1} + d_{13}\alpha_3(\alpha_3 - 1)X^{\alpha_3}) + \\
+ \eta_{12}(d_{31}X^{\alpha_1} + d_{33}X^{\alpha_3} + d_{35}).
\]
For $X_2 > X > X_1$, in the high consumption state:

\[
\rho_1(d_{21}X^{\beta_1} + d_{22}X^{\beta_2} + d_{25} + d_{26}X) = A_0 + \mu_1(d_{21}\beta_1X^{\beta_1} + d_{22}\beta_2X^{\beta_2} + d_{26}X) + \frac{\sigma_1^2}{2}(d_{21}\beta_1(\beta_1 - 1)X^{\beta_1} + d_{22}\beta_2(\beta_2 - 1)X^{\beta_2}) + \eta_{12}lX.
\]

For $X > X_2$, in the low consumptions state:

\[
\rho_2(d_{31}X^{\alpha_1} + d_{33}X^{\alpha_3} + d_{35}) = A_0 + \mu_2(d_{31}\alpha_1X^{\alpha_1} + d_{33}\alpha_3X^{\alpha_3}) + \frac{\sigma_2^2}{2}(d_{31}\alpha_1(\alpha_1 - 1)X^{\alpha_1} + d_{33}\alpha_3(\alpha_3 - 1)X^{\alpha_3}) + \eta_{21}(d_{11}X^{\alpha_1} + d_{13}X^{\alpha_3} + d_{15}).
\]

First consider the case where $X > X_2$. Matching coefficients gives us:

\[
\rho_1d_{15} = A_0 + \eta_{12}d_{35}; \quad (6)
\]

\[
\rho_2d_{35} = A_0 + \eta_{21}d_{15}; \quad (7)
\]

\[
\rho_1d_{11} = \mu_1\alpha_1d_{11} + \frac{\sigma_1^2}{2}\alpha_1(\alpha_1 - 1)d_{11} + \eta_{12}d_{31}; \quad (8)
\]

\[
\rho_2d_{31} = \mu_2\alpha_1d_{31} + \frac{\sigma_2^2}{2}\alpha_1(\alpha_1 - 1)d_{31} + \eta_{21}d_{11}. \quad (9)
\]

From equations (6) and (7),

\[
d_{15} = \frac{(\rho_2 + \eta_{12})}{\rho_1 \rho_2 - \eta_{12} \eta_{21}}A_0;
\]

\[
d_{35} = \frac{(\rho_1 + \eta_{21})}{\rho_1 \rho_2 - \eta_{12} \eta_{21}}A_0.
\]

From equations (8) and (9) we get:

\[
\rho_1 - \mu_1\alpha_1 - \frac{\sigma_1^2}{2}\alpha_1(\alpha_1 - 1) - \eta_{12}d_{11} = \rho_2 - \mu_2\alpha_1 - \frac{\sigma_2^2}{2}\alpha_1(\alpha_1 - 1) - \eta_{21}d_{31}. 
\]
This gives us quadratic equations for $\frac{d_{31}}{d_{11}}$ and $\frac{d_{33}}{d_{13}}$:

\[
\eta_{12}\left(\frac{d_{31}}{d_{11}}\right)^2 + (\rho_2 - \rho_1 + (\mu_1 - \mu_2)\alpha_1 + \left(\frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right)\alpha_1(\alpha_1 - 1))\frac{d_{31}}{d_{11}} - \eta_{21} = 0;
\]

\[
\eta_{12}\left(\frac{d_{33}}{d_{13}}\right)^2 + (\rho_2 - \rho_1 + (\mu_1 - \mu_2)\alpha_3 + \left(\frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right)\alpha_3(\alpha_3 - 1))\frac{d_{33}}{d_{13}} - \eta_{21} = 0.
\]

Define $A_1 = \rho_2 - \rho_1 + (\mu_1 - \mu_2)\alpha_1 + \left(\frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right)\alpha_1(\alpha_1 - 1)$ and $A_3 = \rho_2 - \rho_1 + (\mu_1 - \mu_2)\alpha_3 + \left(\frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right)\alpha_3(\alpha_3 - 1)$. The solutions to the quadratic equations are given by

\[
\lambda_1 \equiv -\frac{A_1 + \sqrt{A_1^2 + 4\eta_{12}\eta_{21}}}{2\eta_{12}} = \frac{d_{31}}{d_{11}};
\]

\[
\lambda_3 \equiv -\frac{A_3 + \sqrt{A_3^2 + 4\eta_{12}\eta_{21}}}{2\eta_{12}} = \frac{d_{33}}{d_{13}}.
\]

The unique fixed points to these equations are $\alpha_1$, $\alpha_3$, $\lambda_1$, and $\lambda_3$ (verify uniqueness).

Now I consider the case where $X_2 > X > X_1$. The ODEs imply the following:

\[
\rho_1 d_{25} = A_0; \quad (10)
\]

\[
\rho_1 d_{26} = \mu_1 d_{26} + \eta_{12} l; \quad (11)
\]

\[
\rho_1 d_{21} = \mu_1 d_{21} \beta_1 + \frac{\sigma_1^2}{2} d_{21} \beta_1 (\beta_1 - 1). \quad (12)
\]

From (10):

\[
d_{25} = \frac{A_0}{\rho_1}.
\]

From (11):

\[
d_{26} = \frac{\eta_{12} l}{\rho_1 - \mu_1}.
\]
From (12):

\[ \beta_1 = \frac{\left( \sigma^2_1 - \mu_1 \right) - \sqrt{\left( \mu_1 - \frac{\sigma^2_1}{2} \right)^2 + 2\sigma^2_1 \rho_1}}{\sigma^2_1}; \]

\[ \beta_2 = \frac{\left( \sigma^2_1 - \mu_1 \right) + \sqrt{\left( \mu_1 - \frac{\sigma^2_1}{2} \right)^2 - 2\sigma^2_1 \rho_1}}{\sigma^2_1}. \]

There are now 6 remaining unknowns: \( d_{11}, d_{31}, d_{21}, d_{22}, d_{31}, d_{33} \), as well as the endogenous thresholds \( \bar{X}_1 \) and \( \bar{X}_2 \). There are 6 independent equations that are linear in the thresholds: the four boundary conditions and the two \( \lambda \) equations. In principle I could solve for the unknown coefficients as a function of the thresholds in closed form. However, since I will need to solve for the thresholds numerically anyway, it is easier to solve these equations numerically at the same time as I solve for the thresholds in the equity problem. The six equations are:

\[ D(\bar{X}_1, S_1) = l\bar{X}_1; \]
\[ D(\bar{X}_2, S_2) = l\bar{X}_2; \]
\[ d_{11}\bar{X}_2^{\alpha_1} + d_{13}\bar{X}_2^{\alpha_3} + d_{15} = d_{21}\bar{X}_2^{\beta_1} + d_{22}\bar{X}_2^{\beta_2} + d_{25} + d_{26}\bar{X}_2; \]
\[ d_{11}\alpha_1\bar{X}_2^{\alpha_1} + d_{13}\alpha_3\bar{X}_2^{\alpha_3} = d_{21}\beta_1\bar{X}_2^{\beta_1} + d_{22}\beta_2\bar{X}_2^{\beta_2} + d_{26}\bar{X}_2; \]
\[ d_{31} = \lambda_1 d_{11}; \]
\[ d_{33} = \lambda_3 d_{13}. \]

1.A.2 Equity Problem

I now solve the equity value problem. Given that the debt value is a polynomial and enters the equity value linearly, it is reasonable to conjecture that the equity value will also be a polynomial. Conjecture

\[ E(X, S_1) = \begin{cases} 
  e_{11}X^{\alpha_1} + e_{13}X^{\alpha_3} + e_{15} + e_{16}X + e_{17}X^{\alpha_7} + e_{19}X^{\alpha_9} & \text{if } X > \bar{X}_2; \\
  e_{21}X^{\beta_1} + e_{22}X^{\beta_2} + e_{25} + e_{26}X + e_{27}X^{\beta_7} + e_{28}X^{\beta_8} & \text{if } \bar{X}_2 > X > \bar{X}_1;
\end{cases} \]
\[ E(X, S_2) = \begin{cases} 
    e_{31}X^{\alpha_1} + e_{33}X^{\alpha_3} + e_{35} + e_{36}X + e_{37}X^{\alpha_7} + e_{39}X^{\alpha_9} & \text{if } X > \bar{X}_2; \\
    0 & \text{if } \bar{X}_2 > X > \bar{X}_1. 
\end{cases} \]

Write

\[
\rho_1^e = r_1 + \eta_{12}; \\
\rho_2^e = r_2 + \eta_{21}; \\
B_1 = -(1 - \pi)c - \phi p + \phi d_{15}; \\
B_2 = -(1 - \pi)c - \phi p + \phi d_{35}; \\
B_3 = -(1 - \pi)c - \phi p + \phi d_{25}; \\
C_1 = \rho_1^e - \mu_1\alpha_1 - \frac{\sigma_1^2}{2}(\alpha_1 - 1); \\
C_2 = \rho_2^e - \mu_2\alpha_1 - \frac{\sigma_1^2}{2}(\alpha_1 - 1); \\
C_3 = \rho_1^e - \mu_1\alpha_3 - \frac{\sigma_2^2}{2}(\alpha_3 - 1); \\
C_4 = \rho_2^e - \mu_2\alpha_3 - \frac{\sigma_2^2}{2}(\alpha_3 - 1). 
\]

The equity value ODEs then imply the following. For \( X > \bar{X}_2 \), in the high consumptions state:

\[
\rho_1^e(e_{11}X^{\alpha_1} + e_{13}X^{\alpha_3} + e_{15} + e_{16}X + e_{17}X^{\alpha_7} + e_{19}x^{\alpha_9}) = B_1 + \delta X + \\
+ \mu_1(e_{11}\alpha_1X^{\alpha_1} + e_{13}\alpha_3X^{\alpha_3} + e_{16}X + e_{17}\alpha_7X^{\alpha_7} + e_{19}\alpha_9x^{\alpha_9}) + \\
+ \frac{\sigma_1^2}{2}(e_{11}\alpha_1(\alpha_1 - 1)X^{\alpha_1} + e_{13}\alpha_3(\alpha_3 - 1)X^{\alpha_3} + e_{17}\alpha_7(\alpha_7 - 1)X^{\alpha_7} + e_{19}\alpha_9(\alpha_9 - 1)x^{\alpha_9}) + \\
+ \eta_{12}(e_{31}X^{\alpha_1} + e_{33}X^{\alpha_3} + e_{35} + e_{36}X + e_{37}X^{\alpha_7} + e_{39}X^{\alpha_9}) + \phi d_{11}X^{\alpha_1} + \phi d_{13}X^{\alpha_3}. 
\]
For $X_2 > X > X_1$, in the high consumption state:

$$\rho_1^e(e_{21}X^{\beta_1} + e_{22}X^{\beta_2} + e_{25} + e_{26}X + e_{27}X^{\beta_7} + e_{28}X^{\beta_8}) = B_3 + \delta X +$$

$$+ \mu_1(e_{21}\beta_1 X^{\beta_1} + e_{22}\beta_2 X^{\beta_2} + e_{26}X + e_{27}\beta_7 X^{\beta_7} + e_{28}\beta_8 X^{\beta_8}) +$$

$$+ \frac{\sigma_1^2}{2}(e_{21}\beta_1(\beta_1 - 1)X^{\beta_1} + e_{22}\beta_2(\beta_2 - 1)X^{\beta_2} + e_{27}\beta_7(\beta_7 - 1)X^{\beta_7} + e_{28}\beta_8(\beta_8 - 1)X^{\beta_8}) +$$

$$+ \phi d_{21}X^{\beta_1} + \phi d_{22}X^{\beta_2} + \phi d_{26}X.$$

For $X > X_2$, in the low consumption state:

$$\rho_2^e(e_{31}X^{\alpha_1} + e_{33}X^{\alpha_3} + e_{35} + e_{36}X + e_{37}X^{\alpha_7} + e_{39}x^{\alpha_9}) = B_2 + \delta X +$$

$$+ \mu_2(e_{31}\alpha_1 X^{\alpha_1} + e_{33}\alpha_3 X^{\alpha_3} + e_{36}X + e_{37}\alpha_7 X^{\alpha_7} + e_{39}\alpha_9 x^{\alpha_9}) +$$

$$+ \frac{\sigma_2^2}{2}(e_{31}\alpha_1(\alpha_1 - 1)X^{\alpha_1} + e_{33}\alpha_3(\alpha_3 - 1)X^{\alpha_3} + e_{37}\alpha_7(\alpha_7 - 1)X^{\alpha_7} + e_{39}\alpha_9(\alpha_9 - 1)x^{\alpha_9}) +$$

$$+ \eta_{21}(e_{11}X^{\alpha_1} + e_{13}X^{\alpha_3} + e_{15} + e_{16}X + e_{17}X^{\alpha_7} + e_{19}X^{\alpha_9}) + \phi d_{31}X^{\alpha_1} + \phi d_{33}X^{\alpha_3}.$$
Consider the case where $X > \bar{X}_2$. Matching coefficients gives us:

$$\rho^e_1 e_{15} = B_1 + \eta_{12} e_{35};$$  \hfill (13)  
$$\rho^e_2 e_{35} = B_2 + \eta_{21} e_{15};$$  \hfill (14)  
$$\rho^e_1 e_{16} = \delta + \mu_1 e_{16} + \eta_{12} e_{36};$$  \hfill (15)  
$$\rho^e_2 e_{36} = \delta + \mu_2 e_{36} + \eta_{21} e_{16};$$  \hfill (16)  

$$\rho^e_1 e_{11} = \mu_1 \alpha_1 e_{11} + \frac{\sigma_1^2}{2} \alpha_1 (\alpha_1 - 1) e_{11} + \eta_{12} e_{31} + \phi d_{11};$$  \hfill (17)  
$$\rho^e_2 e_{31} = \mu_2 \alpha_1 e_{31} + \frac{\sigma_2^2}{2} \alpha_1 (\alpha_1 - 1) e_{31} + \eta_{21} e_{11} + \phi d_{31};$$  \hfill (18)  
$$\rho^e_1 e_{13} = \mu_1 \alpha_3 e_{13} + \frac{\sigma_1^2}{2} \alpha_3 (\alpha_3 - 1) e_{13} + \eta_{12} e_{33} + \phi d_{13};$$  \hfill (19)  
$$\rho^e_2 e_{33} = \mu_2 \alpha_3 e_{33} + \frac{\sigma_2^2}{2} \alpha_3 (\alpha_3 - 1) e_{33} + \eta_{21} e_{13} + \phi d_{33};$$  \hfill (20)  

$$\rho^e_1 e_{17} = \mu_1 \alpha_7 e_{17} + \frac{\sigma_1^2}{2} \alpha_7 (\alpha_7 - 1) e_{17} + \eta_{12} e_{37};$$  \hfill (21)  
$$\rho^e_2 e_{37} = \mu_2 \alpha_7 e_{37} + \frac{\sigma_2^2}{2} \alpha_7 (\alpha_7 - 1) e_{37} + \eta_{21} e_{17};$$  \hfill (22)  
$$\rho^e_1 e_{19} = \mu_1 \alpha_9 e_{19} + \frac{\sigma_1^2}{2} \alpha_9 (\alpha_9 - 1) e_{19} + \eta_{12} e_{39};$$  \hfill (23)  
$$\rho^e_2 e_{39} = \mu_2 \alpha_9 e_{39} + \frac{\sigma_2^2}{2} \alpha_9 (\alpha_9 - 1) e_{39} + \eta_{21} e_{19}. $$  \hfill (24)  

From equations (13) and (14) we get:

$$e_{15} = \frac{B_1 \rho^e_2 + \eta_{12} B_2}{\rho^e_1 \rho^e_2 - \eta_{12} \eta_{21}};$$  
$$e_{35} = \frac{B_2 \rho^e_2 + \eta_{21} B_1}{\rho^e_1 \rho^e_2 - \eta_{12} \eta_{21}}. $$

From equations (15) and (16) we get:

$$e_{16} = \delta \left( \frac{\rho^e_2 - \mu_2 + \eta_{12}}{(\rho^e_1 - \mu_1)(\rho^e_2 - \mu_2) - \eta_{12} \eta_{21}} \right);$$
$$e_{36} = \delta \left( \frac{\rho^e_1 - \mu_1 + \eta_{21}}{(\rho^e_1 - \mu_1)(\rho^e_2 - \mu_2) - \eta_{12} \eta_{21}} \right).$$
Let

\[ F_1 = \frac{\phi(\lambda_1 \eta_2 + C_2)}{C_1 C_2 - \eta_1 \eta_2}; \]
\[ F_2 = \frac{\phi(\frac{\eta_2}{\lambda_1} + C_1)}{C_1 C_2 - \eta_1 \eta_2}; \]
\[ F_3 = \frac{\phi(\lambda_1 \eta_2 + C_4)}{C_3 C_4 - \eta_1 \eta_2}; \]
\[ F_4 = \frac{\phi(\frac{\eta_2}{\lambda_2} + C_3)}{C_3 C_4 - \eta_1 \eta_2}. \]

From equations (17) and (18):

\[ e_{11} = \frac{\phi(\eta_1 d_{31} + C_2 d_{11})}{C_1 C_2 - \eta_1 \eta_2} = F_1 d_{11}; \]
\[ e_{31} = \frac{\phi(\eta_2 d_{11} + C_1 d_{31})}{C_1 C_2 - \eta_1 \eta_2} = F_2 \lambda_1 d_{11}. \]

From equations (19) and (20):

\[ e_{13} = \frac{\phi(\eta_2 d_{33} + C_4 d_{13})}{C_3 C_4 - \eta_1 \eta_2} = F_3 d_{13}; \]
\[ e_{33} = \frac{\phi(\eta_1 d_{21} + C_3 d_{33})}{C_3 C_4 - \eta_1 \eta_2} = F_4 \lambda_3 d_{13}. \]

From equations (21), (22), (23) and (24):

\[ \rho_1^e - \eta_1 \frac{e_{37}}{e_{17}} - \mu_1 \alpha_7 - \frac{\sigma_2}{2} \alpha_7 (\alpha_7 - 1) = \rho_2^e - \eta_2 \frac{e_{17}}{e_{37}} - \mu_2 \alpha_7 - \frac{\sigma_2}{2} \alpha_7 (\alpha_7 - 1); \]
\[ \rho_1^e - \eta_1 \frac{e_{39}}{e_{19}} - \mu_1 \alpha_9 - \frac{\sigma_2}{2} \alpha_9 (\alpha_9 - 1) = \rho_2^e - \eta_2 \frac{e_{19}}{e_{39}} - \mu_2 \alpha_9 - \frac{\sigma_2}{2} \alpha_9 (\alpha_9 - 1). \]

The resulting quadratic equations are:

\[ \eta_1 \left( \frac{e_{37}}{e_{17}} \right)^2 + (\rho_2^e - \rho_1^e + (\mu_1 - \mu_2) \alpha_7 + (\frac{\sigma_2}{2} - \frac{\sigma_2}{2}) \alpha_7 (\alpha_7 - 1)) \left( \frac{e_{37}}{e_{17}} \right) - \eta_2 = 0; \]
\[ \eta_1 \left( \frac{e_{39}}{e_{19}} \right)^2 + (\rho_2^e - \rho_1^e + (\mu_1 - \mu_2) \alpha_9 + (\frac{\sigma_2}{2} - \frac{\sigma_2}{2}) \alpha_9 (\alpha_9 - 1)) \left( \frac{e_{39}}{e_{19}} \right) - \eta_2 = 0. \]
Define $B_7 = \rho_2^e - \rho_1^e + (\mu_1 - \mu_2)\alpha_7 + (\frac{\sigma_2^2}{2} - \frac{\sigma_1^2}{2})\alpha_7(\alpha_7 - 1)$ and $B_9 = \rho_2^e - \rho_1^e + (\mu_1 - \mu_2)\alpha_9 + (\frac{\sigma_2^2}{2} - \frac{\sigma_1^2}{2})\alpha_9(\alpha_9 - 1)$. The solutions to the quadratic equations are given by:

$$
\lambda_1^e \equiv -\frac{B_7 + \sqrt{B_7^2 + 4\eta_{12}\eta_2}}{2\eta_{12}} = e_{37} e_{17};
$$

$$
\lambda_3^e \equiv -\frac{B_9 - \sqrt{B_9^2 + 4\eta_{12}\eta_2}}{2\eta_{12}} = e_{39} e_{19}.
$$

Substituting $e_{37} = \lambda_1^e e_{17}$ and $e_{39} = \lambda_3^e e_{19}$ into equations (21) and (23) gives us two more quadratic equations in $\alpha_7$ and $\alpha_9$:

$$
\alpha_7 = \frac{(\frac{\sigma_1^2}{2} - \mu_1) - \sqrt{(\mu_1 - \frac{\sigma_1^2}{2})^2 - 2\sigma_1^2(\eta_{12}\lambda_1^e - \rho_1^e)}}{\sigma_1^2};
$$

$$
\alpha_9 = \frac{(\frac{\sigma_1^2}{2} - \mu_1) - \sqrt{(\mu_1 - \frac{\sigma_1^2}{2})^2 - 2\sigma_1^2(\eta_{12}\lambda_3^e - \rho_1^e)}}{\sigma_1^2}.
$$

The unique fixed points these equations are $\alpha_7$, $\alpha_9$, $\lambda_1^e$, and $\lambda_3^e$.

I now consider the case where $\bar{X}_2 > X > \bar{X}_1$:

$$
\rho_1^e e_{25} = B_3 l;
$$

$$
\rho_1^e e_{26} = \delta + \mu_1 e_{26} + \phi d_{26};
$$

$$
\rho_1^e e_{21} = \mu_1 \beta_1 e_{21} + \frac{\sigma_1^2}{2} \beta_1(\beta_1 - 1)e_{21} + \phi d_{21};
$$

$$
\rho_1^e e_{22} = \mu_1 \beta_2 e_{22} + \frac{\sigma_1^2}{2} \beta_2(\beta_2 - 1)e_{22} + \phi d_{22};
$$

$$
\rho_1^e e_{27} = \mu_1 \beta_7 e_{27} + \frac{\sigma_1^2}{2} \beta_7(\beta_7 - 1)e_{27};
$$

$$
\rho_1^e e_{28} = \mu_1 \beta_8 e_{28} + \frac{\sigma_1^2}{2} \beta_8(\beta_8 - 1)e_{28}.
$$

41
Equations (25) and (26) imply:

\[ e_{25} = \frac{B_3}{\rho_1^e}; \]
\[ e_{26} = \delta + \phi d_{26} \rho_1^e \left( \frac{\beta_1}{\rho_1^e - \mu_1} \right). \]

Let

\[ C_5 = \frac{\phi}{\rho_1^e - \mu_1 \beta_1 - \frac{\sigma_1^2}{2} \beta_1 (\beta_1 - 1)}; \]
\[ C_6 = \frac{\phi}{\rho_1^e - \mu_1 \beta_2 - \frac{\sigma_1^2}{2} \beta_2 (\beta_2 - 1)}. \]

From equations (27) and (28):

\[ e_{21} = \frac{\phi d_{21}}{\rho_1^e - \mu_1 \beta_1 - \frac{\sigma_1^2}{2} \beta_1 (\beta_1 - 1)} = C_5 d_{21}; \]
\[ e_{22} = \frac{\phi d_{22}}{\rho_1^e - \mu_1 \beta_2 - \frac{\sigma_1^2}{2} \beta_2 (\beta_2 - 1)} = C_6 d_{22}. \]

From (29) and (30):

\[ \beta_7 = \frac{\left( \frac{\sigma_1^2}{2} - \mu_1 \right) - \sqrt{\left( \mu_1 - \frac{\sigma_1^2}{2} \right)^2 + 2 \sigma_1^2 \rho_1^e}}{\sigma_1^2}; \]
\[ \beta_8 = \frac{\left( \frac{\sigma_1^2}{2} - \mu_1 \right) + \sqrt{\left( \mu_1 - \frac{\sigma_1^2}{2} \right)^2 + 2 \sigma_1^2 \rho_1^e}}{\sigma_1^2}. \]

Six unknown equity coefficients remain: \( e_{27}, e_{28}, e_{17}, e_{37}, e_{19}, e_{39} \). In addition, we have the unknown debt coefficients \( d_{11}, d_{31}, d_{21}, d_{22}, d_{31}, d_{33} \), and the endogenous default thresholds \( X_2 \) and \( X_1 \). We have 14 non-linear equations: the four debt boundary equations, the two debt \( \lambda \) equations, the four equity boundary equations, the two equity \( \lambda^e \) equations, and the two equity
smooth pasting conditions:

\[
\begin{align*}
D(X_1, S_1) &= lX_1; \\
D(X_2, S_2) &= lX_2; \\
d_{11}X_2^{\alpha_1} + d_{13}X_2^{\alpha_3} + d_{15} &= d_{21}X_2^{\beta_1} + d_{22}X_2^{\beta_2} + d_{25} + d_{26}X_2; \\
d_{11}\alpha_1 X_2^{\alpha_1} + d_{13}\alpha_3 X_2^{\alpha_3} &= d_{21}\beta_1 X_2^{\beta_1} + d_{22}\beta_2 X_2^{\beta_2} + d_{26}X_2; \\
d_{31} &= \lambda_1 d_{11}; \\
d_{33} &= \lambda_3 d_{13}; \\
E(X_1, S_1) &= 0; \\
E(X_2, S_2) &= 0; \\
\text{VM at } X_2; \\
\text{SP at } X_2; \\
e_{37} &= \lambda_1 e_{17}; \\
e_{39} &= \lambda_3 e_{19}; \\
E_X(X_1, S_1) &= 0; \\
E_X(X_2, S_2) &= 0.
\end{align*}
\]

While we could in principle solve for the coefficients in closed form and then solve for the thresholds numerically, it is simpler to solve the entire system numerically using any standard nonlinear equation solver.
1.B Kalman Filter Details

This appendix explains the Kalman filter approach used to estimate the consumption-liquidity regimes. The algorithm employed is an application of the model produced by Hamilton (1989) and extended by Kim (1994). I implement it in MATLAB.

1.B.1 Notation

I first describe the various vectors and matrices in the model.

- $Y \in \mathbb{M}(3, 1, 96)$ is the matrix of realized signals, giving one 3x1 vector $(\text{Amihud}_t, \text{Roll}_t, S\&P500)'$ for each month $t$ in the 96 months in our sample;
- $L \in \mathbb{M}(2, 1, 96)$ is the unobserved underlying liquidity and consumption series;
- $A \in \mathbb{M}(2, 2, 2)$ is a matrix of autoregressive parameters, where $A(\cdot, \cdot, k)$ is the matrix of autoregressive parameters for the underlying processes $L$ in state $k$;
- $D \in \mathbb{M}(2, 1, 2)$ is a parameter matrix, where $D(i, 1, k)$ is the long-run average of true process $L_i$ in state $k$;
- $F \in \mathbb{M}(3, 2, 2)$ is a parameter matrix that describes how signals $Y$ are produced from the underlying processes $L$: $F(\cdot, \cdot, k)$ is the (expected) mapping from $L$ to $Y$ in state $k$;
- $G \in \mathbb{M}(2, 2, 2)$ is a parameter matrix mapping error terms into observations;
- $N \in \mathbb{M}(2, 2)$ is the Markov transition matrix for the states;
- $R \in \mathbb{M}(3, 3)$ is the covariance matrix for the signal noise $\nu$;
- $Q \in \mathbb{M}(2, 2)$ is the covariance matrix for error term $\epsilon$ in the underlying processes $L$;
- $\hat{L}_{\text{prior}} \in \mathbb{M}(2, 1, 2, 96)$ is a matrix of prior conditional expectations, such that $\hat{L}_{\text{prior}}(\cdot, \cdot, i, j, t)$ is the expected value of $L(\cdot, \cdot, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t - 1$;
• $\hat{L}_{\text{post}} \in \mathbb{M}(2, 1, 2, 2, 96)$ is a matrix of posterior conditional expectations, such that $\hat{L}_{\text{post}}(:, i, j, t)$ is the expected value of $L(:, :, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t$;

• $V_{\text{prior}} \in \mathbb{M}(2, 2, 2, 2, 96)$ is a set of conditional covariance matrices, such that $V_{\text{prior}}(:, i, j, t)$ is the estimated covariance matrix for $\hat{L}_{\text{prior}}(:, i, j, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t - 1$;

• $V_{\text{post}} \in \mathbb{M}(2, 2, 2, 2, 96)$ is a set of posterior conditional covariance matrices, such that $V_{\text{post}}(:, i, j, t)$ is the estimated covariance matrix for $\hat{L}_{\text{post}}(:, i, j, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t$;

• $U \in \mathbb{M}(2, 1, 2, 2, 96)$ is the conditional forecast error matrix, i.e., is $U(:, :, i, j, t)$ is the forecast error of $L(:, :, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t - 1$;

• $H \in \mathbb{M}(2, 2, 2, 2, 96)$ is the conditional variance of forecast matrix, such that $H(:, :, i, j, t)$ is the variance of $U(:, :, i, j, t)$, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t - 1$;

• $K \in \mathbb{M}(2, 2, 2, 2, 96)$ is the Kalman gain matrix, such that $K(:, :, i, j, t)$ is the Kalman gain, conditional on $S_{t-1} = i$, $S_t = j$, and information up to $t - 1$;

• $\Pi_{\text{prior}} \in \mathbb{M}(2, 2, 96)$ is a prior conditional probability matrix, such that $\Pi_{\text{prior}}(i, j, t)$ is the probability that $S_{t-1} = i$ and $S_t = j$, conditional on information up to $t - 1$;

• $\Pi_{\text{post}} \in \mathbb{M}(2, 2, 96)$ is a posterior conditional probability matrix, such that $\Pi_{\text{post}}(i, j, t)$ is the probability that $S_{t-1} = i$ and $S_t = j$, conditional on information up to $t$;

• $\bar{\Pi} \in \mathbb{M}(2, 1, 96)$ is the unconditional probability matrix, such that $\bar{\Pi}(i, 1, t)$ is the probability that $S_t = i$. 

45
1.B.2 Estimation

The algorithm contains four steps. First, I estimate the model using the standard Kalman filter, at each step assuming that $S_{t-1} = i$ and $S_t = j$, for $i, j = 1, 2$. This produces the state-conditional estimates of $L$, $\hat{L}^{ij}$, and the state-conditional estimates of regime probabilities, $\Pi^{ij}$. Second, I collapse these conditional estimates to unconditional estimates, using the conditional probabilities and the Markov transition rates. Third, I estimate the parameters using maximum likelihood. And fourth, I apply the Kalman smoother, giving us period-by-period estimates of the underlying liquidity and consumption processes and the probability of a stress regime using past, present, and future information.

Forming conditional estimates of the underlying process $L$ takes seven steps, for each $t \in 1, \ldots, 96$, and each $i, j \in 1, 2$:

1. Predicting the underlying process from last-periods information:

$$\hat{L}_{\text{prior}}(:, i, j, t) = A(:, j)\hat{L}_{\text{post}}(:, i, j, t-1) + (I - A(:, j))D(:, t);$$

2. Predicting the covariance matrices:

$$V_{\text{prior}}(:, i, j, t) = A(:, j)V_{\text{post}}(:, i, j, t-1)A(:, j)' + G(:, i, j, t)QG(:, i, j, t)';$$

3. Calculating the forecast error:

$$U(:, i, j, t) = Y(:, t) - F(:, j)\hat{L}_{\text{prior}}(:, i, j, t);$$

4. Calculating the variance of the forecast error:

$$H(:, i, j, t) = F(:, j)V_{\text{prior}}(:, i, j, t)F(:, j)' + R;$$
5. Calculating the optimal Kalman gain:

\[ K(:, i, j, t) = V_{\text{prior}}(:, i, j, t)F(:, i, j)H(:, i, j, t)' \]

6. Forming posterior expectations about the underlying processes:

\[ \hat{L}_{\text{post}}(:, i, j, t) = \hat{L}_{\text{prior}}(:, i, j, t) + K(:, i, j, t)U(:, i, j, t) \]

7. Forming the posterior covariance matrix:

\[ V_{\text{post}}(:, i, j, t) = (I - K(:, i, j, t)F(:, i, j))V_{\text{prior}}(:, i, j, t) \]

Likewise, forming conditional estimates of regime probabilities takes three steps:

1. Formulate prior probabilities:

\[ \Pi_{\text{prior}}(i, j, t) = N(j, i)(\Pi_{\text{post}}(1, i, t - 1) + \Pi_{\text{post}}(2, i, t - 1)) \]

2. Formulate conditional state densities:

\[ f(i, j, t) = (2\pi)^{-\frac{1}{2}}|H(:, i, j, t)|^{-\frac{1}{2}}\exp(-0.5U(:, i, j, t)'H(:, i, j, t)^{-1}U(:, i, j, t)) \]

3. Update probabilities:

\[ \Pi_{\text{post}}(i, j, t) = \frac{f(i, j, t)}{\sum_{i} \sum_{j} f(i, j, t)} \]

I then form unconditional probabilities:

\[ \Pi(j, 1, t) = \Pi_{\text{prior}}(1, j, t + 1) + \Pi_{\text{prior}}(2, j, t + 1) \]
I then form log-likelihoods and then maximize them with respect to the parameters:

\[ LL = -\sum_t \log(\sum_i \sum_j f(i, j, t)). \]

The optimal parameters are presented below.

\[
A = \begin{pmatrix}
0.431 & 0 \\
0 & -0.025
\end{pmatrix}, \quad
D = \begin{pmatrix}
0.276 & 1.315 \\
1.508 & -2.014
\end{pmatrix};
\]

\[
F = \begin{pmatrix}
0.656 & 0 \\
0.344 & 0 \\
0 & 1
\end{pmatrix}, \quad
G = \begin{pmatrix}
0.136 & 0.069 \\
0.069 & 0.215 \\
0.104 & 0.179
\end{pmatrix};
\]

\[
N = \begin{pmatrix}
0.962 & 0.038 \\
0.095 & 0.905
\end{pmatrix}; \quad
R = \begin{pmatrix}
0.215 & 0.133 & -0.014 \\
0.133 & 0.179 & -0.008 \\
-0.014 & -0.008 & 0.002
\end{pmatrix}.
\]

The final step is to apply the Kalman smoother, which effectively works backwards from the last date to get smoothed predictions about the state probabilities. Let \( \Pi_{\text{smooth}} \in (2, 2, 95) \) denote the smoothed conditional state probability matrix, such that \( \Pi_{\text{smooth}}(j, k, t) \) is the probability that \( S_{t-1} = j \) and \( S_t = k \), given all information, and let \( \bar{\Pi}_{\text{smooth}}(2, 1, 95) \) be denote the smooth unconditional state probability matrix, such that \( \bar{\Pi}_{\text{smooth}}(i, 1, t) \) is the probability that \( S_t = i \), given the full set. The smoother works backwards using the following algorithm: for \( t = 1, \ldots, 94 \) and \( j, k = 1, 2 \),

\[
\Pi_{\text{smooth}}(j, k, 96 - t) = \frac{\bar{\Pi}_{\text{smooth}}(k, 1, 96 - t)\Pi(j, 1, 95 - t)N(k, j)}{\bar{\Pi}(1, 1, 95 - t)N(k, 1) + \bar{\Pi}(2, 1, 95 - t)N(k, 2)};
\]

\[
\bar{\Pi}_{\text{smooth}}(j, 1, 95 - t) = \Pi_{\text{smooth}}(j, 1, 96 - t) + \Pi_{\text{smooth}}(j, 2, 96 - t).
\]

Figure 2 plots the smoothed month-by-month probabilities of a crisis regime, that is, the series \( \bar{\Pi}_{\text{smooth}}(2, 1, t) \) for \( t = 1, \ldots, 95 \).
<table>
<thead>
<tr>
<th></th>
<th>Spread</th>
<th>Age</th>
<th>Tenor</th>
<th>Issuance Size</th>
<th>Monthly Volume</th>
<th>Daily Volume</th>
<th>Turnover</th>
<th>Amihud Liquidity</th>
<th>Roll Liquidity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td>3.25</td>
<td>5.81</td>
<td>1061</td>
<td>84</td>
<td>4.36</td>
<td>0.12</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>[186]</td>
<td>[2.83]</td>
<td>[4.12]</td>
<td>[751]</td>
<td>[132]</td>
<td>[7.03]</td>
<td>[3.61]</td>
<td>[0.43]</td>
<td>[0.43]</td>
</tr>
<tr>
<td><strong>Aaa and Aa Bonds (23%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>3.10</td>
<td>5.05</td>
<td>1179</td>
<td>80</td>
<td>4.11</td>
<td>0.07</td>
<td>0.30</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>[122]</td>
<td>[2.63]</td>
<td>[3.50]</td>
<td>[761]</td>
<td>[103]</td>
<td>[5.73]</td>
<td>[0.07]</td>
<td>[0.39]</td>
<td>[0.35]</td>
</tr>
<tr>
<td><strong>A Bonds (53%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>145</td>
<td>3.42</td>
<td>5.84</td>
<td>1003</td>
<td>77</td>
<td>3.96</td>
<td>0.07</td>
<td>0.42</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>[161]</td>
<td>[2.91]</td>
<td>[4.15]</td>
<td>[719]</td>
<td>[126]</td>
<td>[6.67]</td>
<td>[0.08]</td>
<td>[0.44]</td>
<td>[0.40]</td>
</tr>
<tr>
<td><strong>Baa Bonds (20%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>251</td>
<td>3.31</td>
<td>6.64</td>
<td>1006</td>
<td>96</td>
<td>5.05</td>
<td>0.25</td>
<td>0.52</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>[247]</td>
<td>[2.95]</td>
<td>[4.52]</td>
<td>[634]</td>
<td>[157]</td>
<td>[8.14]</td>
<td>[7.11]</td>
<td>[0.46]</td>
<td>[0.50]</td>
</tr>
<tr>
<td><strong>Callable Bonds (45%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td>2.93</td>
<td>6.63</td>
<td>990</td>
<td>77</td>
<td>4.12</td>
<td>0.16</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>[158]</td>
<td>[2.47]</td>
<td>[4.30]</td>
<td>[642]</td>
<td>[105]</td>
<td>[6.13]</td>
<td>[5.20]</td>
<td>[0.45]</td>
<td>[0.42]</td>
</tr>
</tbody>
</table>

Table 1: Bond Summary Statistics. The sample includes 1792 distinct bonds and a 96-month period, for a total of 68730 bond-month observations. Means and standard deviations are reported. Spread is the difference between bond yield and the Treasury with the matching month and maturity. Age is the number of years since issuance. Tenor is the number of years remaining until maturity. Issuance size is the total face value of bonds issued, in millions of dollars. Monthly volume is monthly volume traded, in millions of dollars. Daily volume is average daily volume traded, in millions of dollars. Turnover is a bond’s monthly trading volume divided by its issuance size. Amihud is a bond’s Amihud liquidity measure. Roll is a bond’s Roll liquidity measure.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amihud</td>
<td>11.29***</td>
<td>-75.14***</td>
<td>18.98**</td>
<td>-38.41**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.75)</td>
<td>(-3.48)</td>
<td>(3.04)</td>
<td>(-2.73)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44.55***</td>
<td>108.6***</td>
<td>56.39**</td>
<td>82.49**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.99)</td>
<td>(4.05)</td>
<td>(3.26)</td>
<td>(3.43)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.330</td>
<td>-0.865</td>
<td>-1.212</td>
<td>-0.509</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.34)</td>
<td>(-0.71)</td>
<td>(-1.12)</td>
<td>(-0.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maturity</td>
<td>4.683**</td>
<td>4.082**</td>
<td>1.923</td>
<td>1.864</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.74)</td>
<td>(3.37)</td>
<td>(0.98)</td>
<td>(0.93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Issuance</td>
<td>-4.843</td>
<td>-0.685</td>
<td>4.798</td>
<td>0.393</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.35)</td>
<td>(-0.15)</td>
<td>(0.85)</td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Monthly Volume</td>
<td>9.293***</td>
<td>11.56***</td>
<td>14.75***</td>
<td>12.97***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.50)</td>
<td>(4.12)</td>
<td>(4.22)</td>
<td>(4.18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond Beta</td>
<td>45.19**</td>
<td>43.32**</td>
<td>45.06***</td>
<td>45.39***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.33)</td>
<td>(3.21)</td>
<td>(3.51)</td>
<td>(3.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Callable dummy</td>
<td>-58.47*</td>
<td>-58.25*</td>
<td>-52.24*</td>
<td>-49.42*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.52)</td>
<td>(-2.50)</td>
<td>(-2.60)</td>
<td>(-2.60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial dummy</td>
<td>12.12</td>
<td>11.17</td>
<td>8.296</td>
<td>8.604</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.43)</td>
<td>(1.38)</td>
<td>(1.12)</td>
<td>(1.18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A dummy</td>
<td>33.65**</td>
<td>28.77**</td>
<td>31.04**</td>
<td>53.96**</td>
<td>49.71**</td>
<td>49.08**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.28)</td>
<td>(3.24)</td>
<td>(3.28)</td>
<td>(2.70)</td>
<td>(2.76)</td>
<td>(2.78)</td>
<td></td>
</tr>
<tr>
<td>Baa dummy</td>
<td>111.5***</td>
<td>99.59***</td>
<td>99.61***</td>
<td>132.7***</td>
<td>119.2***</td>
<td>118.3***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.67)</td>
<td>(4.88)</td>
<td>(4.82)</td>
<td>(3.42)</td>
<td>(3.71)</td>
<td>(3.77)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>103.6***</td>
<td>94.75***</td>
<td>96.90***</td>
<td>112.4**</td>
<td>72.69***</td>
<td>20.51</td>
<td>59.29***</td>
</tr>
<tr>
<td></td>
<td>(4.12)</td>
<td>(4.21)</td>
<td>(4.35)</td>
<td>(2.92)</td>
<td>(3.84)</td>
<td>(1.30)</td>
<td>(4.04)</td>
</tr>
<tr>
<td>Average $R^2$</td>
<td>0.170</td>
<td>0.202</td>
<td>0.227</td>
<td>0.321</td>
<td>0.326</td>
<td>0.339</td>
<td>0.345</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Spreads and Liquidity. Monthly Fama-Macbeth regressions with bond spread as the dependent variable. $t$-stats are corrected for possible serial correlation using Newey-West standard errors and are reported in parentheses. *Amihud* and *Roll* are the Amihud and Roll liquidity measured constructed using the full sample. *Age, Tenor, Log Issuance, Log Monthly Volume,* and *Bond Beta* are as described in Table 1. Average $R^2$ is the time-series average of the cross-sectional $R^2$s.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-Crisis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amihud</td>
<td>18.98**</td>
<td>25.63**</td>
<td>19.13**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.04)</td>
<td>(2.77)</td>
<td>(3.44)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roll</td>
<td>56.39**</td>
<td>42.91**</td>
<td>18.90**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.26)</td>
<td>(2.87)</td>
<td>(3.67)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crisis Amihud</td>
<td>45.21*</td>
<td>51.17***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.27)</td>
<td>(4.49)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crisis Roll</td>
<td></td>
<td>23.40**</td>
<td></td>
<td>63.78**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.77)</td>
<td>(3.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.865</td>
<td>-1.212</td>
<td>-2.156*</td>
<td>-1.756</td>
<td>-3.389*</td>
<td>-2.504*</td>
</tr>
<tr>
<td></td>
<td>(-0.71)</td>
<td>(-1.12)</td>
<td>(-2.12)</td>
<td>(-2.45)</td>
<td>(-2.98)</td>
<td>(-2.44)</td>
</tr>
<tr>
<td>Tenor</td>
<td>4.082**</td>
<td>1.923</td>
<td>2.561</td>
<td>0.788</td>
<td>5.984***</td>
<td>3.522</td>
</tr>
<tr>
<td></td>
<td>(3.37)</td>
<td>(0.98)</td>
<td>(1.93)</td>
<td>(0.41)</td>
<td>(3.81)</td>
<td>(1.57)</td>
</tr>
<tr>
<td>Log Issuance</td>
<td>-0.685</td>
<td>4.798</td>
<td>-0.562</td>
<td>5.947</td>
<td>-17.58**</td>
<td>-8.117**</td>
</tr>
<tr>
<td></td>
<td>(-0.15)</td>
<td>(0.85)</td>
<td>(-0.16)</td>
<td>(1.35)</td>
<td>(-3.54)</td>
<td>(-3.08)</td>
</tr>
<tr>
<td></td>
<td>(4.12)</td>
<td>(4.22)</td>
<td>(3.34)</td>
<td>(3.23)</td>
<td>(7.21)</td>
<td>(7.31)</td>
</tr>
<tr>
<td>Bond Beta</td>
<td>43.32**</td>
<td>45.06***</td>
<td>29.52**</td>
<td>31.86**</td>
<td>28.63**</td>
<td>31.96***</td>
</tr>
<tr>
<td></td>
<td>(3.21)</td>
<td>(3.51)</td>
<td>(3.18)</td>
<td>(3.50)</td>
<td>(11.08)</td>
<td>(11.34)</td>
</tr>
<tr>
<td>Callable dummy</td>
<td>-58.25*</td>
<td>-52.24*</td>
<td>-55.38*</td>
<td>-49.56*</td>
<td>-82.42***</td>
<td>-75.53***</td>
</tr>
<tr>
<td></td>
<td>(-2.50)</td>
<td>(-2.60)</td>
<td>(-2.40)</td>
<td>(-2.41)</td>
<td>(-4.09)</td>
<td>(-4.25)</td>
</tr>
<tr>
<td>Financial dummy</td>
<td>11.17</td>
<td>8.296</td>
<td>10.71</td>
<td>7.570</td>
<td>25.29</td>
<td>22.32</td>
</tr>
<tr>
<td></td>
<td>(1.38)</td>
<td>(1.12)</td>
<td>(1.62)</td>
<td>(1.25)</td>
<td>(1.65)</td>
<td>(1.77)</td>
</tr>
<tr>
<td>A dummy</td>
<td>53.30**</td>
<td>49.71**</td>
<td>51.28**</td>
<td>47.35**</td>
<td>60.75**</td>
<td>54.42**</td>
</tr>
<tr>
<td></td>
<td>(2.71)</td>
<td>(2.76)</td>
<td>(2.64)</td>
<td>(2.70)</td>
<td>(3.10)</td>
<td>(3.01)</td>
</tr>
<tr>
<td>Baa dummy</td>
<td>130.3***</td>
<td>119.2***</td>
<td>113.7**</td>
<td>105.9**</td>
<td>133.3**</td>
<td>119.4**</td>
</tr>
<tr>
<td></td>
<td>(3.50)</td>
<td>(3.71)</td>
<td>(3.05)</td>
<td>(3.20)</td>
<td>(3.29)</td>
<td>(3.27)</td>
</tr>
<tr>
<td>Constant</td>
<td>72.69***</td>
<td>20.51**</td>
<td>72.90*</td>
<td>17.58</td>
<td>126.5**</td>
<td>40.71**</td>
</tr>
<tr>
<td></td>
<td>(3.84)</td>
<td>(1.30)</td>
<td>(2.56)</td>
<td>(1.74)</td>
<td>(3.36)</td>
<td>(2.60)</td>
</tr>
<tr>
<td>Average R²</td>
<td>0.326</td>
<td>0.339</td>
<td>0.378</td>
<td>0.391</td>
<td>0.292</td>
<td>0.311</td>
</tr>
</tbody>
</table>

Table 3: Spreads and Crisis Liquidity. Monthly Fama-Macbeth regressions with bond spread as the dependent variable. t-stats are corrected for possible serial correlation using Newey-West standard errors and are reported in parentheses. Columns (1) through (4) use the full sample of spreads, while columns (5) and (6) use only spreads from the post-crisis period (May 2009 to March 2011). *Amihud* and *Roll* are the Amihud and Roll liquidity measured using the full sample. *Crisis Amihud* and *Crisis Roll* are the Amihud and Roll liquidity measures constructed using observations between August 2007 and April 2009. *Age, Tenor, Log Issuance, Log Monthly Volume,* and *Bond Beta* are as described in Table 1. Average R² is the time-series average of the cross-sectional R²s.
Figure 1: Macroeconomic variables and the crisis. Panel (a) plots monthly aggregate illiquidity under the Amihud measure. Panel (b) plots monthly aggregate illiquidity under the Roll measure. Panel (c) plots the equally-weighted spread average of investment-grade bonds in our sample. Panel (d) plots the monthly S&P500 total return.
Figure 2: Monthly stress regime probabilities implied by maximum likelihood estimation of the Markov-state Kalman filter. For details, see Appendix 1.B.
Figure 3: Calibrated Model Performance, full sample (1781 bonds). Model predicted spreads versus actual spreads, in (a) the regime-varying liquidity model, using the Amihud measure of liquidity; (b) the regime-varying liquidity model, using the Roll measure of liquidity; (c) the constant liquidity model, using the Amihud measure of liquidity; (d) the constant liquidity model, using the Roll measure of liquidity.
Figure 4: Calibrated Model Performance, A-rated bonds (784 bonds). Model predicted spreads versus actual spreads, in (a) the regime-varying liquidity model, using the Amihud measure of liquidity; (b) the regime-varying liquidity model, using the Roll measure of liquidity; (c) the constant liquidity model, using the Amihud measure of liquidity; (d) the constant liquidity model, using the Roll measure of liquidity.
Figure 5: Calibrated Model Performance, Pre- and Post- Crisis. Panel (a) plots actual pre-crisis spreads versus spreads predicted by the constant-liquidity baseline model. Panel (b) plots actual pre-crisis spreads versus spreads predicted by the regime-varying liquidity model. Panel (c) plots post-crisis spreads versus spreads predicted by the constant-liquidity model. Panel (d) plots post-crisis spreads versus spreads predicted by the regime-varying model. All panels use A-rated bonds only and the Roll illiquidity measure.
2 On-the-run and Off-the-run Corporate Bond Liquidity

2.1 Introduction

Any study of asset liquidity premia faces two immediate challenges. First, how does one measure asset liquidity? And second, how does one identify the premium in the asset’s price or rate of return that is associated with liquidity? The standard approach is to posit one or more measures of liquidity, and then measure the associated premium using panel regressions. A variety of liquidity measures have been developed for this purpose, such as the bid-ask spread (Amihud and Mendelson (1986)), persistent price impact (Amihud (2002) and Acharya and Pedersen (2005)), transitory price impact (Pastor and Stambaugh (2003)), or transactions costs implied by a history of traded prices (Roll (1984), Bao et al. (2010), Edwards et al. (2007), and Chen et al. (2007)).

In this paper, I investigate the liquidity and liquidity premia of corporate bonds using a unique identification method. The essence of this method is the so-called on-the-run effect: for a given bond issuer, the “on-the-run bond”, namely, the bond that was most recently issued, is anecdotally more liquid than other “off-the-run” bonds. This effect is widely accepted in the market for US Treasuries. Moreover, in the market for Treasuries, this liquidity effect results in a premium. For example, using daily data between 1995 and 1999, Krishnamurthy (2002) shows that the on-the-run 30-year Treasury bond has an average yield of 6.25 basis points lower than the most recent off-the-run 30-year Treasury. Similarly, Fleming (2003) shows that the average spread between the yields of on-the-run and off-the-run 10-year Treasury notes is 5.63 basis points, between 1996 and 2000. To my knowledge, this is the first paper that explicitly compares on-the-run and off-the-run corporate bonds.

My empirical approach combines Mergent’s comprehensive Fixed Income Securities

7In Robinson (2012) I provide a comprehensive review of the empirical and theoretical literature on asset liquidity and liquidity premia.
Database with FINRA’s Transaction Reporting and Compliance Engine and identifies
a set of 504 large, liquid bonds belonging to 36 distinct issuers. Then, using issue and
transaction data from 2002 to 2011, I measure liquidity for each bond at a monthly level,
and relate bond liquidity to characteristics such as age, rating, industry, and in partic-
ular, whether the bond is on-the-run. I find strong evidence that on-the-run corporate
bonds are more liquid than comparable off-the-run bonds. My final endeavor is to de-
termine whether off-the-run bonds trade at higher yields than their on-the-run counter-
parts. To investigate this I construct monthly portfolios that hold one long position in
each (most-recent) off-the-run bond, and one short position in each on-the-run bond.
The yield spread on this portfolio is positive but only marginally significant. More in-
teresting is the observation that the spread explodes during the crisis month of October
2008, suggesting that the spread is possibly related to liquidity.

The rest of the paper is organized as follows. In Section 2.2, I describe my data sets
and construct liquidity measures. In Section 2.3, I investigate the determinants of liq-
uidity. In particular, I address the question of whether on-the-run corporate bonds are
more liquid than comparable off-the-run bonds. In Section 2.4, I turn to the question of
whether on-the-run bonds command a liquidity premium. More specifically: does a port-
folio of off-the-run bonds command a yield spread over a portfolio of on-the-run bonds,
and if so, what are the historical properties of this spread? In Section 2.5 I conclude.

2.2 Data Description

I employ two corporate bond data sets for my analysis of corporate bond liquidity. The
first is FINRA’s Transaction Reporting and Compliance Engine (TRACE), which con-
tains data on over 100 million bond transactions beginning in 2002, including date of
transaction, time of transaction, par volume traded, clean price, and yield. To obtain
bond-specific characteristics, such as the bond issuer name and industry, the dollar size
of the issue, dates of issuance and maturity, ratings, and features such callability, putability, convertibility, and coupon type, I merge the TRACE data with Mergent’s Fixed Income Securities Database (FISD), a comprehensive corporate bond database with over 200,000 issues.

2.2.1 Bonds, Issuers, and Transactions

My focus is on corporate bonds that are most directly comparable to Treasury notes; it is for these bonds that the on-the-run/off-the-run liquidity differential will be most relevant and identifiable. This means bonds that are large, liquid, vanilla, and relatively credit-worthy. With this in mind, I remove short-term bonds (bonds with a maturity of less than five years at issuance) and bonds with exotic features such as putability and convertibility, perpetual bonds, and bonds with floating coupons. I keep callable bonds, which comprise almost 50 percent of the sample. Because a large part of my analysis will involve comparing multiple bonds from the same issuer, I drop all bonds whose issuers have fewer than nine bonds in the sample. Tables 4 and 5 summarize my final sample, which includes 504 bonds from 36 distinct issuers such as Citigroup, General Electric Capital, and IBM. Using North American Industry Classification System (NAICS) codes, I classify each issuer as financial or non-financial; 11 out of 36 (30%) of the issuers are classified as financial, and they issue 37% of the bonds. Finally, using the Mergent FISD ratings sub-dataset, I generate a time-series of ratings for each bond. Table 5 shows the median bond rating for each issuer at issuance; the median rating overall may be lower for bonds issuers that are downgraded midway through the sample (ie, Lehman).

With these bonds in hand I proceed to the TRACE transaction data, which is matched with FISD using bond-specific CUSIPs. Each TRACE observation contains the date

---

8Distinguishing between financials and non-financials will be important in much of my analysis, in light of the fact that my sample period includes the 2007-2009 financial crisis and my issuer sample includes Bear Sterns and Lehman Brothers.
and time of the transaction, the volume traded (in par value), the clean price (per hundred dollars of par value), the yield to maturity, and whether the transaction was inter-dealer. A few features of the data are readily apparent. First, there is substantial variation in transaction size: the minimum transaction recorded is $1000 par value, while the maximum transaction recorded is $5MM+. Second, the distribution of transaction sizes is heavily skewed: the median transaction size is only $25 thousand, while the average transaction size is $314 thousand. The standard interpretation of this feature is that there are broadly two types of transactions happening in the corporate bond market: small “retail” transactions, and large “institutional” transactions. Because these two types of transactions potentially behave very differently, I (somewhat arbitrarily) label trades of less than $50,000 in par value as “retail” trades, and trades of greater than $50,000 par value as “institutional” trades, which will allow me to conduct separate analyses. Finally, it is important to mention that 6% of the transactions in the sample are truncated; namely, they are over $5 million (for investment trade bonds) or $1 million (for junk bonds) in par value, and are therefore recorded as “5MM+” or “1MM+” in TRACE respectively. I replace the volumes of these trades with 5 million and 1 million respectively, but I also label these trades as “large” trades (as well as “institutional” trades), for separate analysis.

The transaction data is summarized in Panel B of Table 5. The size, frequency, and number of transactions for these 504 bonds is consistent with them being an extremely liquid subsample of all corporate bonds. After eliminating duplicate inter-dealer transactions and transactions with incorrectly entered data, I am left with approximately 8 million transactions, or about 15,700 transactions per bond, over the 114 months between July 2002 and December 2011. Bonds are traded on an average of 1,036 days (out of a total of 2394 trading days in the sample period), and are traded an average of 15.4 times per day on days that they are traded at least once. The average daily par volume traded
per bond is $4.8 million. Finally, approximately 54% of transactions are classified as “re-
tail” (less than $50 thousand in par value was traded), while 46% are classified as “in-
stitutional” (over $50 thousand in par value), and 6% are “large” (over $5 million or $1
million in par value for investment grade and junk bonds, respectively).

2.2.2 Measuring Liquidity

I now turn to the construction of liquidity measures. I choose four measures for liquid-
ity: number of transactions, volume traded, turnover, and finally, the “Roll” bid-ask
bounce measure. The first three measures are consistent with the premise that is it eas-
ier to trade, and easier to trade large quantities of, a liquid security. The final measure,
the Roll “bid-ask bounce” is consistent with the premise that a more liquid security will
entail a smaller transaction cost. Additionally, each of these four measures can be con-
structed with the full sample of transactions, as well as the sub-samples that include only
retail transactions, institutional transactions, or large transactions.

The core of my analysis will involve monthly liquidity measures, under the assump-
tion that a bond’s liquidity is constant over each month. I thus collapse the 8 million
transaction data set into a data set containing 26,980 bond-month observations. Each
bond-month observation contains the following monthly liquidity measures: total par
value traded that month (as well as the par value of retail and institutional trades); to-
tal number of transactions (as well as the number of retail, institutional, and large trans-
actions); turnover that month, defined by the ratio of par value traded to issuance size
(also retail and institutional turnover); and finally, the monthly Roll illiquidity measure,
constructing using all transactions, as well as just institutional transactions. Appendix
2.A describes how the Roll illiquidity measure is constructed.

Table 6 summarizes the four bond-month liquidity measures for the whole sample,
for financial bonds only (which comprise about 40% of bond-months), for junk bonds
only (10%), and for bonds that are currently “on-the-run” (13% of the sample). A cursory examination of the median statistics suggests that financial bonds are somewhat more liquid than non-financials (at least, as measured by par volume traded and average number of transactions); for example, the average monthly volume traded for financial bonds is $102 million, while the average monthly volume traded for all bonds is $90 million. Along these lines, junk bonds are less liquid than investment grades bonds ($57m compared to $90m). And finally, on-the-run bonds are more liquid than off-the-run bonds ($179m, compared to $90m). Of course, this cursory analysis does not control for anything; a detailed analysis of the determinants of bond liquidity is presented in the following section.

2.3 Determinants of Bond Liquidity

In this section I examine the determinants of bond liquidity, using the sample of 26,980 bond-months and our four measures of liquidity. The ultimate goal will be to determine if on-the-run bonds are more liquid than off-the-run bonds, controlling for other bond and economic variables. However, before examining the on-run effect, I look at two other potential determinants of bond liquidity: age, and time (time series variation in aggregate liquidity).

2.3.1 Bond Age and Liquidity

A number of studies have shown that bond liquidity decreases as the bond ages. For example, Hotchkiss and Jostova (2007) and Alexander et al. (2000) find a strong negative correlation between bond age and trading volume; Bao et al. (2010) find that Roll illiquidity is positively correlated with age; Edwards et al. (2007) find that implied transaction costs increase with age; and Chen et al. (2007) find that age is positively correlated with a variety of illiquidity measures. One anecdotal explanation for this relationship
is that a substantial fraction of bond portfolios are “inactive”; once bonds “settle” into these portfolios, they are traded very infrequently.

I perform a simple age-liquidity analysis by sorting the 26,980 bond-months into buckets by age in months. At this stage of the analysis, I do not control for other potentially important variables, such as issuance size, rating, and whether the bond is on-the-run or not. I measure liquidity using trading volume, number of transactions, monthly turnover, and Roll illiquidity; each measure is further subdivided to include all transactions, retail transactions only, institutional transactions only, and large transactions only. Figure 6 plots average monthly par volume traded, in millions of dollars, against bond age, in months, over the first 24 months of each bond’s life. Panel (a) shows that there is a clear and unmistakable association between age and trading volume: for example, the average par volume traded for a one-month-old bond is $300 million, whereas average par volume traded for a twelve-month-old bond is only $100 million. Panels (b) and (c) decompose the trading volume into retail and institutional transactions. As expected, the vast majority of volume at every level is due to institutional transactions. However, these panels also reveal an interesting difference between the trading behavior of retail and institutional traders. Whereas institutional volume monotonically decreases with age, retail volume remains relatively constant as the bond ages, after an initial decrease in volume following the first two months of trading. Possible explanations for this are discussed below.

Figure 8 plots monthly turnover, defined by the ratio of par volume traded to issuance size, against age. The results qualitatively mirror those of Figure 6, showing the results discussed in the previous paragraph are not driven by large issuances. On average, a bond experiences a 25% turnover in the month following its offering. One year later, its monthly turnover falls to an average of 8%. Retail and institutional turnover share the same features as retail and institutional volume: average institutional turnover
monotonically decreases with age, while average retail turnover falls sharply for two months, then remains relatively constant as the bond ages.

Figure 7 plots the average number of monthly transactions per bond against age. Three observations deserve mentioning. First, unlike volume, transactions are divided relatively evenly between the retail and institutional classes; in fact, retail transactions make up slightly more than half of the transaction sample overall. Second, as with volume, bonds experience a large number of transactions in the first few months following issuance. However, unlike volume, the number of transactions per month stabilizes relatively quickly (after 3-4 months) at around 250-300 transactions per month. And third, like with volume, institutional transactions experience a near-monotonic decline as the bond ages, while retail transactions experience no trend after a sharp fall following the first two months after issuance.

Finally, I examine how the Roll illiquidity measure varies with age. Figure 9 plots average monthly Roll illiquidity against age. For each bond-month, I construct Roll illiquidity using the method described in Appendix 2.A; bond-months are then pooled and averaged by age. Figure 9 shows a slight upward trend, suggesting that bonds are becoming more illiquid as they age. However, there is quite a lot of variation around the trend. I conduct a more careful analysis in Section 2.3.3.

To summarize, the analysis of this section demonstrates that bonds become more illiquid as they age, an effect that is strongest in the first few months following issuance. This result holds for a variety of liquidity measures and transaction types. Therefore, it will be important to control for bond age in any analysis of whether on-the-run bonds are more liquid than off-the-run bonds.
2.3.2 Time Series Trends in Aggregate Liquidity

I now examine whether there are any time-series trends in aggregate bond liquidity. If such trends exist, we will need to account for them in the analysis of on-the-run liquidity, for example, by adding time fixed effects. One obvious hypothesis is that aggregate bond liquidity decreased during the financial crisis; previous literature (for example, Bao et al. (2010)) has demonstrated this to be true. To construct the aggregate monthly liquidity measures, I start with the 26,980 bond-month observations and take the cross-sectional average of each liquidity measure in each month. The sample period is then truncated to begin in February of 2005. Aggregate liquidity measures before this date are biased for two reasons. First, because TRACE reporting was slowly phased in, beginning in 2002, and starting with the most liquid bonds, the early years of my sample are biased towards liquid bonds. Second, because bonds enter the data set at a faster rate than they leave in the early years of the sample, average bond age is substantially lower between 2002 and 2004, which of course is correlated with higher liquidity. For the most part, these sample biases have disappeared by 2005.

Figure 10 illustrates the time trend in aggregate volume traded. While aggregate total volume traded certainly experiences a lot of month-to-month variation, there is no discernable trend, at least at first glance. This is also true for aggregate institutional volume. In sharp contrast, aggregate retail volume traded experiences a huge increase that begins in late 2008 and lasts until late 2009, coinciding with the financial crisis: over this period, retail volume per bond is between $4 and $8 million per month, compared to $1 to $2 million per month prior, and $2 to $3 million post. Figure 11 plots the time trends in the number of aggregate transactions per bond. Total transactions, retail transactions, and institutional transactions all exhibit stark increases during the crisis period, in contrast to volume. For example, total transactions per bond increase from 100-200 per month pre-crisis, to 300-800 per month during the crisis, and then fall back to 200-
300 per month post-crisis. The aggregate time series results for volume and transactions reveal an interesting phenomenon. These two measures are frequently associated with liquidity *cross-sectionally*; it seems natural to think of a bond as more liquid if it trades more frequently than another, otherwise similar bond. This association is also consistent with our liquidity-age panel analysis; younger bonds are widely considered more liquid and trade more frequently than older bonds. However, the time series analysis shows that aggregate transactions and (retail) volume dramatically increased in a period widely associated with asset *illiquidity*. One implication is that we should be careful to control for time-series effects in a panel analysis of liquidity.

Finally, Figure 12 illustrates the time series trend in aggregate Roll illiquidity over time. Aggregate Roll illiquidity is constructed each month by averaging monthly Roll for each bond, discarding bonds-months that include fewer than 10 transactions. In contrast to the aggregate volume and transaction time series, the Roll time series shows that aggregate liquidity sharply *falls* during the crisis period of mid 2008 to mid 2009. This result mirrors the similar Roll liquidity analyses in Bao et al. (2010) and Robinson (2012).

### 2.3.3 Are On-the-run Bonds More Liquid? A Panel Analysis

The final component of my analysis is a series of panel regressions. The primary goal is to determine whether on-the-run bonds are more liquid than off-the-run bonds, controlling for a variety of other firm, bond, and time characteristics. Once again, the data set will be the 26,980 bond-month observations constructed in the previous section. My re-
gressions are classified into four types of specifications:

\[ L_{i,t} = \beta_i X_{i,t} + \epsilon_{i,t}; \quad \text{(No fixed effects)} \]
\[ L_{i,t} = \beta_i X_{i,t} + \delta_j + \epsilon_{i,t}; \quad \text{(Issuer fixed effects)} \]
\[ L_{i,t} = \beta_i X_{i,t} + \alpha_i + \epsilon_{i,t}; \quad \text{(Bond fixed effects)} \]
\[ L_{i,t} = \beta_i X_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}. \quad \text{(Bond and month fixed effects)} \]

\( L_{i,t} \) is the liquidity of bond \( i \) in month \( t \), as measured by volume, number of transactions, turnover, or (negative) Roll. \( X_{i,t} \) is a vector of bond- and time-specific regressors including age, time-to-maturity, issuance size, rating, a financial dummy, and a dummy for on-the-run. For issuer \( j \), bond \( i \), and month \( t \), \( \delta_j \), \( \alpha_i \), and \( \gamma_t \) are issuer-\-, bond-\-, and time specific fixed effects respectively. And finally, \( \epsilon_{i,t} \) is a mean-zero residual that is uncorrelated with the regressors. All standard errors are clustered at the bond level.

Table 7 presents the output for the first class of regression. Many coefficients are significant. For example, the coefficients on the age regressor imply that increasing a bond’s age by one month decreases its monthly volume traded by an average of $1 million. Interestingly, the coefficients on age for the transactions and Roll measures are positive, though small. Larger issuances are unambiguously more liquid, as shown by the positive and significant coefficients on the issuance size regressor. Similarly, investment grade bonds are also unambiguously more liquid than junk bonds.

Our main interest is in the coefficient for the on-the-run dummy. The hypothesis is that on-the-run bonds are more liquid than off-the-run bonds, controlling for other characteristics. The results of the above regressions broadly support the hypothesis. All three coefficients are significant and positive. For example, the results suggest that an on-the-run bond has monthly trading volume of almost $80 million more than a similar off-the-run bond (the average monthly volume traded is $90 million). Likewise, an on-the-run
bond is traded an average of 68 times per month more than a comparable off-the-run bond (the average bond is traded 290 times per month).

Table 8 presents the output for the regressions that include issuer-specific fixed effects. The assumption behind issuer-specific fixed effects is that all bonds belonging to a given issuer experience similar average levels of liquidity. The results of the issuer fixed effect regressions are qualitatively similar to the previous results. The coefficients on the on-the-run dummy are again positive and significant in the volume and transaction specifications; the coefficient is positive but insignificant in the Roll specification.

Table 9 presents the output for the regression models that include bond fixed effects and both bond and time fixed effects. The assumption behind bond-specific fixed effects is that each bond has a different average level of liquidity. The assumption behind time-specific fixed effects is that aggregate liquidity is time-varying; the results of Section 2.3.2 broadly support this idea. The results of the regressions support the hypothesis that on-the-run bonds are more liquid than off-the-run bonds. All six coefficients for the on-the-run dummy are significant and positive. For example, in the model that includes both time and month fixed effects, the coefficient for transactions is 95.36. The interpretation is that the average on-the-run bond trades 95 times more than it would have if it were off-the-run in the same month.

2.4 Analyzing the Spread between On-the-run and Off-the-run Corporate Bond Yield Spreads

In the previous section, I provided evidence that on-the-run bonds are more liquid than their off-the-run counterparts. Intuitively, market participants should require a higher rate of return, ie a higher yield spread, to hold less liquid bonds. This is indeed the case for Treasury securities: for example, Krishnamurthy (2002) documents that the liquid on-the-run 30-year Treasury bond traded at an average discount of 6.25 basis points compared to the less liquid off-the-run (issued 6 months earlier) 30-year Treasury between
This difference in yield between two otherwise similar securities suggests a trading strategy that capitalizes on a liquidity premium: hold the off-the-run bond, and sell short the on-the-run bond. Under the assumption that both bonds default simultaneously (i.e., neither bond is preferred), this trade should not be exposed to small changes in credit risk or credit risk premia. Indeed, this liquidity trade was made (in)famous by its aggressive implementation by Long Term Capital Management prior to its collapse. As is well known (see, for example, Lowenstein (2001) or Furfine and Remolona (2002)), the spread between the on-the-run yield and the off-the-run yield widened dramatically as the price of liquidity increased, causing huge losses for LTCM and contributing to its demise.

### 2.4.1 Constructing a Spread Trade

In this section I examine the spread on a related trade: one that is long off-the-run and short on-the-run corporate bonds. The strategy is constructed using the following algorithm and the 26,980 bond-months from the previous sections. For each month $t$, the algorithm looks for issuers (out of the 36 total issuers) with at least two bonds currently outstanding; call this set of issuers $I(t)$ with cardinality $n(t)$. It then constructs a portfolio that has a short-position in the $n(t)$ on-the-run bonds, and a long position in the $n(t)$ most recent off-the-run bonds. The trade thus involves exactly $2n(t)$ bonds each month, 2 from each issuer. An important assumption of this trade is that the short positions completely fund the long positions.\(^9\)

I am interested in addressing two specific questions. First, is the spread on this trade significantly different from zero? In other words, is there an on-the-run premium for (my subset of) corporate bonds? And second, how has this spread behaved historically? In-

---

\(^9\)Krishnamurthy (2002) shows that the difference in repo rates between on-the-run and off-the-run Treasuries is large enough to explain the difference in their respective yield-to-maturities.
so far as this spread is a proxy for a liquidity premium, its dynamics might tell us something about the evolution of the price of liquidity between 2003 and 2011. Of course, before addressing these questions, we must first calculate a monthly time series of spreads. There are at least three ways to define the spread on this trade. The first, which replicates the calculation of the spread on on-the-run and off-the-run Treasuries, is to simply define the spread as the difference between the yield on the long portfolio containing \( n(t) \) off-the-run bonds, and the yield on the short portfolio containing \( n(t) \) on-the-run bonds. However, this method misses a potentially important point: the durations of the two portfolios may be quite different. Unlike 30-year Treasuries, which are issued every 6 months, corporate bonds are issued at irregular and less frequent intervals. Indeed, a summary analysis of the on-the-run off-the-run pairs shows that the median difference in time-to-maturity between the on-the-run bond and the off-the-run bond is 11 months. Moreover, 11 months is even more significant once one accounts for the fact that the majority of bonds in the sample are medium term notes (5-10 years) and not 30-year bonds.

Because yield curves are typically upwards sloping, the above spread calculation would likely underestimate a liquidity premium. The remaining two spread calculation methods are designed to compensate for this effect. The second method uses the Treasury yield curve to hedge away this difference in duration, or, to put it another way, involves taking long and short positions in Treasury securities that match the durations of the corporate bond portfolios. For example, suppose that the on-the-run portfolio in month \( t \) containing \( n(t) \) bonds has a duration of \( D^{on}(t) \), while the off-the-run portfolio in month \( t \) (also containing \( n(t) \) bonds) has a duration of \( D^{off}(t) \), with \( D^{on}(t) > D^{off}(t) \). The net duration of the long-short portfolio is thus \( D^{off}(t) - D^{on}(t) < 0 \). Method two hedges the difference in duration by taking a long position in \( n(t) \) Treasuries with duration \( D^{on}(t) \) and a short position in \( n(t) \) Treasuries with duration \( D^{off}(t) \), resulting in a net duration of zero.
Method two accounts for the upwards slope of the Treasury yield curve. However, it fails to account for another stylized fact: that the yield curve of corporate bonds typically has a larger slope than the Treasury yield curve, or in other words, the “credit spread” curve is also upwards sloping. Method three accounts for this by hedging away the difference in duration using corporate bond yield curves. To implement this, I make use of a much larger data set that I constructed in Robinson (2012), containing 1781 investment-grade corporate bonds. This allows me to construct IG corporate bond yield curves for each month in my data set. The hedging method can thus be interpreted as follows. Again supposing the on-the-run position in month $t$ contains $n(t)$ bonds with a duration of $D^{on}(t)$ and the off-the-run position contains $n(t)$ bonds with duration $D^{off}(t)$, the hedging strategy takes a long position in a diversified portfolio of investment grade bonds with duration $D^{on}(t)$, and a short position in a diversified portfolio of investment grade bonds with duration $D^{off}(t)$. The result is again a portfolio with a net duration of zero, but unlike the previous method, this method accounts for the fact that the difference in yields for a given duration difference is likely greater for corporate bonds than it is for Treasuries.

### 2.4.2 The On-the-run Off-the-run Corporate Bond Spread

Panel (a) of Figure 14 plots the monthly spread on the three portfolios constructed using method one, method two, and method three. Four features of the graph are immediately apparent. First, the mean of the spread is close to zero, which is consistent with the spread being quite small (as it is for Treasuries). This is especially true for the period prior to mid-2008. Second, the spread seems to experience mean-reversion. Third, the spread experiences a dramatic spike in the month of October 2008 (the month fol-
lowing the collapse of Lehman Brothers).\textsuperscript{10} And fourth, this spike is then followed by a massive increase in variance; the spread is strongly heteroskedastic.

Our first question of interest is whether the mean of this spread is significantly different from zero. Panel A of Table 10 lists the means, \(t\)-statistics, and significance levels of the spread under the three portfolio construction methods. The mean spreads under the three methods are 2.93 bps, 7.48 bps, and 8.33 bps, which is consistent with the upward-sloping Treasury and corporate yield curves hypothesized above. However, the mean spreads of methods one and two are not statistically different from zero. Moreover, when the standard errors are corrected for auto-correlation and heteroskedasticity using the Newey-West procedure, the mean spread of method 3 (8.33 bps) is also not statistically different from zero.

Our second question of interest is how the spread behaves over the sample period. As mentioned above, the behavior is stark: the spread mean reverts to (approximately) zero with low variance prior to October 2008. It then spikes upwards to almost 200 bps and retains a much higher variance afterwards. Inspection of the individual pairs of bonds in October 2008 reveals that this spike is not due to any single firm (recall that Lehman Brothers was removed from the sample for this analysis), but rather a systemic increase in the spread between the on-the-run and off-the-run bonds of each firm. This is consistent with the idea that this spread is driven by liquidity, and that liquidity experienced a massive drying-up in October 2008, similar to the widening of the Treasury spread that crushed LTCM.

In relating this spread to the price of liquidity, an implicit assumption is that the trade is neutral to credit risk. Intuitively, this seems correct; after all, the trade is long one bond and short one bond of each firm, and preferred securities are excluded. A de-\textsuperscript{10}All Lehman Brothers bonds themselves are excluded from this entire analysis. See below for a discussion.
fault should therefore have no effect on the value of the portfolio. While this is true, it misses a key feature of a bond’s yield: it depends not only on the market value (price) of the bond, but also on its duration. This means that as a firm approaches default, and the prices of its bonds converge to an expected recovery amount, the yields of its bonds diverge, with shorter maturity bonds trading at much higher yields. Figure 15 illustrates this effect for the on-the-run and most recent off-the-run Lehman Brothers bonds in August and September of 2008. The on-the-run bond LEH.JAD is a 7-year bond issued in September of 2007 and maturing in September of 2014; the off-the-run bond LEH.HEO is a 5-year bond issued in July of 2007 and maturing in July of 2012. Throughout August of 2008, the yield spreads on these two bonds remain nearly identical, reflected the premise that these bonds have equal default risk. However, as Lehman approaches bankruptcy in September and the prices of both bonds plummet, the yield spreads diverge by over 10 percentage points, reflecting the fact that as prices converge to low values, yields diverge if durations differ.

Of course, this effect potentially obscures the liquidity effect, which is why I remove Lehman from the sample (if Lehman is included, the spread is substantially wider in October 2008, because the on-the-run bond LEH.JAD has a longer duration and thus a lower spread). This effect may be present to a lesser extent with other issuers. To control for this, I construct portfolios using only high-grade issuers, defined as issuers whose median bond rating is Aa or Aaa; 14 out of the 36 issuers qualify. Panel (b) of Figure 14 plots the method-one, method-two, and method-three spreads of the portfolios containing bonds from only these issuers, and Panel B of Table 10 presents their means, t-statistics, and significance levels. Qualitatively, the results are the same; the spread has a mean of around 7 bps. One important difference is that the spread is now significantly different from zero (at the 5% significance level) for the portfolios constructing using methods two and three.
2.5 Conclusion

In this paper I broadly address two questions. First, are on-the-run bonds more liquid than similar off-the-run bonds? And second, does this difference in liquidity command a premium? In Section 2.3, I present evidence that broadly suggests that the answer to the first question is yes: on-the-run bonds are significantly more liquid than off-the-run bonds, where liquidity is measured in a variety of ways, and issuer, bond, and time controls are included. In Section 2.4, I address the second question by comparing the yields on two portfolios, one which contains each issuer’s on-the-run bond, and one which contains the most recent off-the-run bond from each issuer. The evidence is admittedly weak: while the yield spread is between 5-8 basis points, the significance levels are relatively low. Of greater interest is the historical path of the spread: it explodes to a level of 200 basis points in October of 2008. This is reminiscent of the dramatic widening of the spread between on-the-run and off-the-run Treasury bonds that contributed to the demise of LTCM. And it is consistent with the “flight to liquidity” that anecdotally characterized both episodes.
2.A The Roll Illiquidity Measure

This appendix describes the intuition behind, and the construction of, the Roll (1984) illiquidity measure. This measure is based on the premise that a bond will trade at a higher price if the transaction is buyer-initiated, and a lower price if the transaction is seller-initiated. Under the assumption that the transaction-initiator is independent of the initiator of the previous transaction, we would expect price changes to have negative serial covariance due to this bid-ask bounce, in contrast to a random walk model, where price changes are independent. Moreover, a greater covariance (in absolute terms) implies a larger bid-ask bounce and hence a more illiquid asset. The Roll method thus allows one to construct a measure of liquidity from a time series of transaction prices.

The key assumptions of the Roll measure are as follows:

1. Observed prices $P_t$ consist of two components: a fundamental component $f_t$, and a liquidity component $e_t$. The liquidity component reflects the fact that buy orders transact at a different price than sell orders: $e_t = s/2$ for buy orders, and $e_t = -s/2$ for sell orders, where $s$ is the bid-ask spread.

2. Changes in fundamentals $f_t$ are not predictable, ie, $cov(\Delta f_t, \Delta f_{t+1}) = 0$.

3. Order flow (the arrival of buy orders and sell orders) is not predictable.

Under these assumptions, $cov(\Delta P_t, \Delta P_{t+1}) = cov(\Delta e_t, \Delta e_{t+1})$, since $cov(\Delta f_t, \Delta f_{t+1}) = 0$ by assumption 2 and $cov(\Delta e_t, \Delta e_{t+1}) = 0$ and $cov(\Delta f_t, \Delta e_{t+1}) = 0$ by assumption 3. Moreover, $cov(\Delta e_t, \Delta e_{t+1})$ will be negative because of the switching between buy and sell orders, and this covariance will be increasing (in absolute value) in the bid-ask spread $s$. Thus the negative autocovariance of price changes $-cov(\Delta P_t, \Delta P_{t+1})$ is a measure that is positively correlated with
unobserved transaction costs, and it thus serves as a proxy for bond illiquidity. Therefore, using a time-series of transaction prices for each bond, I construct monthly Roll illiquidity measures by estimating the (negative) autocovariance in price changes. The average bond-month contains 290 transactions; I follow Bao et al. (2010) and require the bond-month to include at least 10 transactions for its Roll measure to be valid. The summary statistics for the monthly Roll measure are presented in Table 6.
### Panel A: Bonds and Issuers

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of bonds</td>
<td>504</td>
</tr>
<tr>
<td>Number of issuers</td>
<td>36</td>
</tr>
<tr>
<td>Percent of bonds issued by financial firms</td>
<td>37%</td>
</tr>
<tr>
<td>Percent of bonds investment grade at issuance</td>
<td>89%</td>
</tr>
<tr>
<td>Average issuance size</td>
<td>$1bn</td>
</tr>
<tr>
<td>Average maturity at issuance</td>
<td>7.6 years</td>
</tr>
<tr>
<td>Trading days in sample</td>
<td>2394</td>
</tr>
</tbody>
</table>

### Panel B: Transactions

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of transactions</td>
<td>7,934,862</td>
</tr>
<tr>
<td>Average number of transactions per bond</td>
<td>15,743</td>
</tr>
<tr>
<td>Median transaction size</td>
<td>$25k</td>
</tr>
<tr>
<td>Average transaction size</td>
<td>$314k</td>
</tr>
<tr>
<td>Average days traded per bond</td>
<td>1,011</td>
</tr>
<tr>
<td>Average daily transactions per bond</td>
<td>15.4</td>
</tr>
<tr>
<td>Average daily par volume traded per bond</td>
<td>$4.8m</td>
</tr>
<tr>
<td>Fraction of transactions that are retail</td>
<td>54%</td>
</tr>
<tr>
<td>Fraction of transactions that are institutional</td>
<td>46%</td>
</tr>
<tr>
<td>Fraction of transactions that are large</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 4: Bond, Issuer, and Transaction Summary Statistics
<table>
<thead>
<tr>
<th>Issuer</th>
<th>Number of Bonds</th>
<th>Median Bond Rating at Issuance</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMERICAN GEN FIN CORP</td>
<td>10</td>
<td>A</td>
</tr>
<tr>
<td>BANK AMER CORP</td>
<td>20</td>
<td>Aa</td>
</tr>
<tr>
<td>BEAR STEARNS COS INC</td>
<td>11</td>
<td>A</td>
</tr>
<tr>
<td>BERKSHIRE HATHAWAY FIN CORP</td>
<td>10</td>
<td>Aaa</td>
</tr>
<tr>
<td>CATERPILLAR FINL SVCS CORP</td>
<td>10</td>
<td>A</td>
</tr>
<tr>
<td>CHESAPEAKE ENERGY CORP</td>
<td>10</td>
<td>Ba</td>
</tr>
<tr>
<td>CIT GROUP INC</td>
<td>20</td>
<td>A</td>
</tr>
<tr>
<td>CITIGROUP INC</td>
<td>31</td>
<td>Aa</td>
</tr>
<tr>
<td>COMCAST CORP</td>
<td>12</td>
<td>Baa</td>
</tr>
<tr>
<td>CREDIT SUISSE FIRST BOSTON (USA) INC</td>
<td>10</td>
<td>Aa</td>
</tr>
<tr>
<td>DOW CHEM CO</td>
<td>11</td>
<td>Baa</td>
</tr>
<tr>
<td>DU PONT E I DE NEMOURS CO</td>
<td>11</td>
<td>A</td>
</tr>
<tr>
<td>FORD MTR CR CO</td>
<td>18</td>
<td>Ba</td>
</tr>
<tr>
<td>GENERAL ELEC CAP CORP</td>
<td>39</td>
<td>Aaa</td>
</tr>
<tr>
<td>GMAC LLC</td>
<td>10</td>
<td>Ca</td>
</tr>
<tr>
<td>GOLDMAN SACHS GROUP INC</td>
<td>15</td>
<td>Aa</td>
</tr>
<tr>
<td>HARRAHS OPER CO INC</td>
<td>11</td>
<td>Baa</td>
</tr>
<tr>
<td>HCA HEALTHCARE CO</td>
<td>11</td>
<td>Ba</td>
</tr>
<tr>
<td>HEWLETT PACKARD CO</td>
<td>12</td>
<td>A</td>
</tr>
<tr>
<td>HOUSEHOLD FIN CORP</td>
<td>17</td>
<td>Aa</td>
</tr>
<tr>
<td>INTERNATIONAL BUSINESS MACHS CORP</td>
<td>10</td>
<td>A</td>
</tr>
<tr>
<td>J P MORGAN CHASE CO</td>
<td>26</td>
<td>Aa</td>
</tr>
<tr>
<td>KRAFT FOODS INC</td>
<td>13</td>
<td>A</td>
</tr>
<tr>
<td>LEHMAN BROS IHDGS INC</td>
<td>15</td>
<td>A</td>
</tr>
<tr>
<td>MERRILL LYNCH CO INC</td>
<td>15</td>
<td>Aa</td>
</tr>
<tr>
<td>MORGAN STANLEY</td>
<td>9</td>
<td>Aa</td>
</tr>
<tr>
<td>PROCTER GAMBLE CO</td>
<td>10</td>
<td>Aa</td>
</tr>
<tr>
<td>SHELL INTL FIN B V</td>
<td>9</td>
<td>Aa</td>
</tr>
<tr>
<td>SIMON PPTY GROUP LP</td>
<td>12</td>
<td>A</td>
</tr>
<tr>
<td>SLM CORP</td>
<td>12</td>
<td>A</td>
</tr>
<tr>
<td>TARGET CORP</td>
<td>9</td>
<td>A</td>
</tr>
<tr>
<td>TIME WARNER CABLE INC</td>
<td>10</td>
<td>Baa</td>
</tr>
<tr>
<td>VODAFONE GROUP PLC</td>
<td>10</td>
<td>A</td>
</tr>
<tr>
<td>WAL MART STORES INC</td>
<td>19</td>
<td>Aa</td>
</tr>
<tr>
<td>WELLS FARGO CO</td>
<td>14</td>
<td>Aa</td>
</tr>
<tr>
<td>XEROX CORP</td>
<td>11</td>
<td>Baa</td>
</tr>
</tbody>
</table>

Table 5: Bond Issuers
<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Financial</th>
<th>Junk</th>
<th>On-the-run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond-month observations</td>
<td>26,980</td>
<td>11,285</td>
<td>2,662</td>
<td>3,448</td>
</tr>
<tr>
<td>Average monthly par volume traded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>90m</td>
<td>102m</td>
<td>57m</td>
<td>179m</td>
</tr>
<tr>
<td>Retail</td>
<td>2.8m</td>
<td>3.2m</td>
<td>2.0m</td>
<td>2.6m</td>
</tr>
<tr>
<td>Institutional</td>
<td>88m</td>
<td>99m</td>
<td>55m</td>
<td>176m</td>
</tr>
<tr>
<td>Average monthly transactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>290</td>
<td>320</td>
<td>245</td>
<td>307</td>
</tr>
<tr>
<td>Retail</td>
<td>183</td>
<td>203</td>
<td>146</td>
<td>164</td>
</tr>
<tr>
<td>Institutional</td>
<td>107</td>
<td>118</td>
<td>99</td>
<td>143</td>
</tr>
<tr>
<td>Large</td>
<td>8.8</td>
<td>7.2</td>
<td>3.1</td>
<td>19.9</td>
</tr>
<tr>
<td>Average monthly turnover (percent)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7.3</td>
<td>7.3</td>
<td>4.9</td>
<td>13.7</td>
</tr>
<tr>
<td>Retail</td>
<td>0.25</td>
<td>0.26</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>Institutional</td>
<td>7.0</td>
<td>7.1</td>
<td>4.8</td>
<td>13.5</td>
</tr>
<tr>
<td>Average monthly Roll illiquidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.41</td>
<td>0.44</td>
<td>0.46</td>
<td>0.42</td>
</tr>
<tr>
<td>Institutional</td>
<td>1.00</td>
<td>1.06</td>
<td>1.51</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table 6: Monthly Bond Liquidity Measures
Table 7: Liquidity Panel Regressions, No Fixed Effects. The unit of observation is a bond-month. I estimate regressions of the form $L_{i,t} = \beta_t X_{i,t} + \epsilon_{i,t}$, where $L_{i,t}$ is the liquidity of bond $i$ in month $t$, $X_{i,t}$ is a set of bond- and month-specific regressors, and $\epsilon_{i,t}$ is a mean-zero residual. Liquidity is measured by (1) monthly volume (in millions), (2) monthly turnover, (3) transactions, and (4) negative Roll. The regressors are bond age (in months), a dummy for on-the-run, bond issuance size, bond years-to-maturity at issuance, a dummy for investment grade, and a dummy for financial issuer. All standard errors are clustered at the bond level. $t$ statistics are reported in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>(1) Volume</th>
<th>(2) Turnover</th>
<th>(3) Transactions</th>
<th>(4) -Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-48.07***</td>
<td>37.42***</td>
<td>-100.9**</td>
<td>-0.0489</td>
</tr>
<tr>
<td></td>
<td>(-4.19)</td>
<td>(6.10)</td>
<td>(-2.65)</td>
<td>(-1.46)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.919***</td>
<td>-0.710***</td>
<td>1.767***</td>
<td>0.00421***</td>
</tr>
<tr>
<td></td>
<td>(-10.93)</td>
<td>(-12.52)</td>
<td>(5.60)</td>
<td>(14.78)</td>
</tr>
<tr>
<td>On-the-run</td>
<td>79.21***</td>
<td>57.20***</td>
<td>67.85**</td>
<td>0.0453*</td>
</tr>
<tr>
<td></td>
<td>(5.99)</td>
<td>(7.99)</td>
<td>(2.65)</td>
<td>(2.27)</td>
</tr>
<tr>
<td>Issuance Size</td>
<td>69.45***</td>
<td>1.486</td>
<td>209.2***</td>
<td>0.00852</td>
</tr>
<tr>
<td></td>
<td>(13.55)</td>
<td>(0.66)</td>
<td>(10.49)</td>
<td>(0.86)</td>
</tr>
<tr>
<td>Maturity at Issuance</td>
<td>5.400***</td>
<td>3.416***</td>
<td>2.838</td>
<td>-0.0727***</td>
</tr>
<tr>
<td></td>
<td>(4.89)</td>
<td>(4.98)</td>
<td>(0.99)</td>
<td>(-18.56)</td>
</tr>
<tr>
<td>Investment Grade</td>
<td>38.89***</td>
<td>27.31***</td>
<td>58.35*</td>
<td>0.0914***</td>
</tr>
<tr>
<td></td>
<td>(6.47)</td>
<td>(8.57)</td>
<td>(2.53)</td>
<td>(3.45)</td>
</tr>
<tr>
<td>Financial</td>
<td>-5.526</td>
<td>-2.030</td>
<td>-26.12</td>
<td>-0.0417**</td>
</tr>
<tr>
<td></td>
<td>(-0.88)</td>
<td>(-0.53)</td>
<td>(-1.25)</td>
<td>(-2.94)</td>
</tr>
</tbody>
</table>

$N$ | 26980 | 26980 | 26980 | 26792 |
$R^2$ | 0.251 | 0.133 | 0.193 | 0.099 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
<table>
<thead>
<tr>
<th></th>
<th>(1) Volume</th>
<th>(2) Turnover</th>
<th>(3) Transactions</th>
<th>(4) -Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-151.4*</td>
<td>-15.76</td>
<td>-211.7*</td>
<td>-0.0109</td>
</tr>
<tr>
<td></td>
<td>(-2.43)</td>
<td>(-0.51)</td>
<td>(-2.14)</td>
<td>(-0.21)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.959***</td>
<td>-0.737***</td>
<td>1.620***</td>
<td>0.00410***</td>
</tr>
<tr>
<td></td>
<td>(-11.14)</td>
<td>(-13.32)</td>
<td>(5.71)</td>
<td>(14.51)</td>
</tr>
<tr>
<td>On-the-run</td>
<td>83.24***</td>
<td>60.24***</td>
<td>94.30***</td>
<td>0.0356</td>
</tr>
<tr>
<td></td>
<td>(6.67)</td>
<td>(8.93)</td>
<td>(4.11)</td>
<td>(1.84)</td>
</tr>
<tr>
<td>Issuance Size</td>
<td>65.36***</td>
<td>-1.447</td>
<td>191.7***</td>
<td>0.00865</td>
</tr>
<tr>
<td></td>
<td>(16.13)</td>
<td>(-0.74)</td>
<td>(11.17)</td>
<td>(0.85)</td>
</tr>
<tr>
<td>Maturity at Issuance</td>
<td>5.736***</td>
<td>3.558***</td>
<td>6.483*</td>
<td>-0.0746***</td>
</tr>
<tr>
<td></td>
<td>(5.21)</td>
<td>(5.27)</td>
<td>(2.30)</td>
<td>(-21.53)</td>
</tr>
<tr>
<td>Investment Grade</td>
<td>161.9*</td>
<td>83.39**</td>
<td>168.9</td>
<td>0.0102</td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(2.59)</td>
<td>(1.72)</td>
<td>(0.18)</td>
</tr>
</tbody>
</table>

| $N$            | 26980      | 26980        | 26980            | 26792     |
| $R^2$          | 0.304      | 0.186        | 0.272            | 0.116     |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Liquidity Panel Regressions, Issuer Fixed Effects. The unit of observation is a bond-month. I estimate regressions of the form $L_{i,t} = \beta_i X_{i,t} + \delta_j + \epsilon_{i,t}$, where $L_{i,t}$ is the liquidity of bond $i$ in month $t$, $X_{i,t}$ is a set of bond- and month-specific regressors, $\delta_j$ is a fixed effect specific to issuer $j$, and $\epsilon_{i,t}$ is a mean-zero residual. Liquidity is measured by (1) monthly volume (in millions), (2) monthly turnover, (3) transactions, and (4) negative Roll. The regressors are bond age (in months), a dummy for on-the-run, bond issuance size, bond years-to-maturity at issuance, and a dummy for investment grade. All standard errors are clustered at the bond level. $t$ statistics are reported in parentheses.
### Panel A: Bond Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>(1) Volume</th>
<th>(2) Turnover</th>
<th>(3) Transactions</th>
<th>(4) -Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>61.65***</td>
<td>66.94***</td>
<td>5.603</td>
<td>-0.508***</td>
</tr>
<tr>
<td></td>
<td>(10.80)</td>
<td>(20.33)</td>
<td>(0.42)</td>
<td>(-30.89)</td>
</tr>
<tr>
<td>Age</td>
<td>-1.115***</td>
<td>-0.868***</td>
<td>1.839</td>
<td>0.00482*</td>
</tr>
<tr>
<td></td>
<td>(-8.13)</td>
<td>(-11.13)</td>
<td>(5.96)</td>
<td>(12.86)</td>
</tr>
<tr>
<td>On-the-run</td>
<td>110.0***</td>
<td>81.29***</td>
<td>112.3***</td>
<td>0.0627*</td>
</tr>
<tr>
<td></td>
<td>(6.52)</td>
<td>(9.25)</td>
<td>(4.37)</td>
<td>(2.57)</td>
</tr>
<tr>
<td>N</td>
<td>26980</td>
<td>26980</td>
<td>26980</td>
<td>26792</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.407</td>
<td>0.296</td>
<td>0.411</td>
<td>0.164</td>
</tr>
</tbody>
</table>

### Panel B: Bond and Month Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>(5) Volume</th>
<th>(6) Turnover</th>
<th>(7) Transactions</th>
<th>(8) -Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>355.0***</td>
<td>194.2***</td>
<td>363.2***</td>
<td>-1.316***</td>
</tr>
<tr>
<td></td>
<td>(4.62)</td>
<td>(5.59)</td>
<td>(3.91)</td>
<td>(-9.72)</td>
</tr>
<tr>
<td>Age</td>
<td>-1.632***</td>
<td>-0.950***</td>
<td>-0.348</td>
<td>0.00512***</td>
</tr>
<tr>
<td></td>
<td>(-11.01)</td>
<td>(-9.95)</td>
<td>(-1.05)</td>
<td>(13.50)</td>
</tr>
<tr>
<td>On-the-run</td>
<td>100.7***</td>
<td>75.49***</td>
<td>95.36***</td>
<td>0.0980***</td>
</tr>
<tr>
<td></td>
<td>(6.36)</td>
<td>(9.06)</td>
<td>(4.36)</td>
<td>(4.59)</td>
</tr>
<tr>
<td>N</td>
<td>26980</td>
<td>26980</td>
<td>26980</td>
<td>26792</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.443</td>
<td>0.343</td>
<td>0.496</td>
<td>0.292</td>
</tr>
</tbody>
</table>

Table 9: Liquidity Panel Regressions, with Bond and Month Fixed Effects. The unit of observation is a bond-month. In Panel A I estimate regressions of the form $L_{i,t} = \beta_i X_{i,t} + \alpha_i + \epsilon_{i,t}$, where $L_{i,t}$ is the liquidity of bond $i$ in month $t$, $X_{i,t}$ is a set of bond- and month-specific regressors, $\alpha_i$ is a fixed effect for bond $i$, and $\epsilon_{i,t}$ is a mean-zero residual. The regressors are bond age and a dummy for on-the-run. In Panel B, I estimate regressions of the form $L_{i,t} = \beta_i X_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}$, where $\gamma_t$ is a fixed effect for month $t$. All standard errors are clustered at the bond level. $t$ statistics are reported in parentheses.
<table>
<thead>
<tr>
<th>Panel A: All issuers</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (bps)</td>
<td>2.93</td>
<td>7.48</td>
<td>8.33</td>
</tr>
<tr>
<td>t-stat</td>
<td>0.76</td>
<td>1.92</td>
<td>2.20*</td>
</tr>
<tr>
<td>Newey-West t-stat</td>
<td>0.62</td>
<td>1.55</td>
<td>1.79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: High-grade issuers</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (bps)</td>
<td>-2.35</td>
<td>7.10</td>
<td>7.55</td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.63</td>
<td>2.41*</td>
<td>2.54*</td>
</tr>
<tr>
<td>Newey-West t-stat</td>
<td>-0.52</td>
<td>1.99*</td>
<td>2.03*</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 10: The On-the-run Off-the-run Corporate Bond Spread Trade. Details of the construction of the spread trade are described in Section 2.4. Panel A uses all available bonds (from 36 issuers) to construct the trade. Panel B uses only issuers whose median bond rating is at least Aa (14 issuers). Method 1 calculates the spread as the difference in portfolio yields. Method 2 hedges differences in duration between the on-the-run and off-the-run portfolios using the Treasury yield curve. Method 3 hedges difference in duration using an aggregate investment-grade corporate bond yield curve.
Figure 6: Monthly Par Volume Traded by Age. Average monthly par volume (in millions of dollars) by bond age (in months) for (a) all transactions, (b) transactions of less than $50 thousand in par value, and (c) transactions of more than $50 thousand in par value.
Figure 7: Monthly Transactions by Age. Average monthly transactions by age (in months): (a) all transactions, (b) transactions of less than $50 thousand in par value, (c) transactions of more than $50 thousand in par value, and (d) transactions of more than $5 million (or $1 million for junk bonds) in par value.
Figure 8: Monthly Turnover by Age. Average monthly turnover (measured by par volume traded, divided by issuance size) by age (in months): (a) all transactions, (b) transactions of less than $50 thousand in par value, and (c) transactions of more than $50 thousand in par value.
Figure 9: Monthly Roll Illiquidity by Age. Average monthly Roll illiquidity, for (a) all transactions, and (b) transactions of more than $50 thousand in par value. Only bond-months with at least 10 transactions of the given type are included. See Appendix 2.A for details on the construction of the Roll measure.
Figure 10: Aggregate Monthly Volume per Bond by Month. Each month, an aggregate monthly volume per bond measure is constructed by taking the average of par volume traded (a) in all transactions, (b) in retail transactions, and (c) in institutional transactions.
Figure 11: Aggregate Monthly Transactions per Bond, by Month. Each month, an aggregate monthly transactions per bond measure is constructed by taking the average number of transactions over all bonds traded, where (a) all transactions are counted, (b) only retail transactions are counted, (c) only institutional transactions are counted, and (d) only large transactions are counted.
Figure 12: Aggregate Monthly Roll Illiquidity, by Month. Each month, Roll illiquidity is constructed for each bond, using the method detailed in Appendix 2.A. Aggregate monthly Roll is constructed by averaging the Roll measures over all bonds, each month. Only bond-months with at least 10 transactions are included. Panel (b) plots the Roll measure that is constructed using only institutional trades.
Figure 13: Average Yield Spreads by Month. Panel (a) plots the average yield spread (calculated using closing prices) of the bonds in the sample each month. Panel (b) plots the average yield spreads of the investment grade bonds (dashed line) and the junk grade bonds (solid line) each month. Panel (c) plots the average yield spreads of the non-financial bonds (dashed line) and financial bonds (solid line) each month. All spreads are calculated as the difference between the bond yield and the yield on a US Treasury of comparable duration.
Figure 14: Spreads on the On-the-run / Off-the-run Long-Short Trades, by Month. The long and short portfolios are constructed by identifying the on-the-run and most recent off-the-run bonds for each issuer and month in the sample, and taking a short and long position in each respectively. Panel (a) uses the full sample of bonds, and calculates the spread using three methods. With method one, duration is not hedged. With method two, duration is hedged using the Treasury yield curve. With method three, duration is hedged using an aggregate investment-grade corporate bond yield curve. Panel (b) repeats the analysis using only issuers whose median bond ratings are Aa or Aaa.
Figure 15: Lehman On-the-run and Off-the-run Price and Yield Spreads, August-September 2008. Panel (a) plots the daily closing prices of the most recently issued Lehman Brothers bond LEH.JAD and the second-most-recently-issued LEH.HEO during the months of August and September 2008. Panel (b) plots their yield spreads. The yields visibly diverge as the prices fall to default levels.
3 Non-linear Hedge Fund Risk Exposures

3.1 Introduction

Since Sharpe (1992), investors and researchers have sought to identify the risk factors correlated with a given manager’s investment strategy. Identifying underlying risk factors is valuable since it allows investors to analyze how a manager’s returns correlate with other investments in the investor’s portfolio, how the manager is likely to perform in out-of-sample scenarios, and whether the manager provides “excess” returns adjusting for risk. For these reasons, many researchers have extended Sharpe (1992)’s initial work to the analysis of hedge fund returns; for example, see Schneeweis and Spurgin (1998), Liang (1999), Edwards and Caglayan (2001), Ennis and Sebastian (2003), Capocci and Hubner (2004), Kat and Palaro (2006), and Hasanhodzic and Lo (2007). The standard approach is to perform linear regressions of the returns of individual hedge funds or hedge fund indices on a variety of market risk factors.

A key insight is that many hedge funds deliver returns that are non-linear functions of commonly used risk-factors such as market equity returns (Fung and Hsieh 1997, 2001; Mitchell and Pulvino 2001). As a result, risk factor models that fail to take these non-linearities into account, such as the factors based on Sharpe (1964), Lintner (1965), and Fama and French (1993) may significantly mis-characterize risk exposure, especially in extreme market situations. Glosten and Jagannathan (1994) propose the use of options as risk factors to capture these non-linearities. Many researchers have used the returns on options or a portfolio of options to determine the risk factors to which hedge funds are exposed (Agarwal and Naik 2004; Fung and Hsieh 2004; Jurek and Stafford 2011). In particular, Agarwal and Naik (2004) estimate risk factor models to explain the returns on Hedge Fund Research’s hedge fund indices between 1994 and 2000 using standard risk factors as well as returns on S&P500 index options.
In this paper, we extend the analysis of Agarwal and Naik (2004) in two ways. First, we expand the time period analyzed by ten years (through March 2011) so that our analysis includes the 2008 financial crisis. Our hope is that the presence of additional downside tail events in our sample (the financial crisis, as well as the dot-com bubble) will yield additional insights about non-linear return relationships. Second, we perform an additional analysis using Hedge Fund Research’s daily hedge fund indices, which started in 2003. Using daily returns dramatically increases the number of sample observations, which hopefully also allows us to see more tail events.

Using monthly returns, we are able to replicate Agarwal and Naik (2004)’s finding that the returns on many hedge fund strategies are negatively correlated with at-the-money or close to at-the-money S&P500 put options prior to 2000. When we extend the sample through 2011, we find that the hedge fund strategies are no longer correlated with these put strategies. However, visual inspection of the hedge fund index returns suggests put-like nonlinearities with respect to the S&P 500 return, and we find that the strategies are correlated with deeper out-of-the-money puts. Likewise, using the daily data, we find that three out of five hedge fund indices experience daily return correlations with out-of-the-money puts.

The remainder of the paper is organized as follows. Section 3.2 describes the hedge fund returns and risk factors used for the monthly analysis. Section 3.3 presents the model and results of the monthly analysis. Section 3.4 describes the data used for the daily analysis. Section 3.5 presents the results from the daily analysis. Section 3.6 concludes.
3.2 Monthly Data Sources

3.2.1 Hedge Fund Indices

We proxy hedge fund returns using Hedge Fund Research’s Hedge Fund Indices (HFR Indices from here on). These indices are not investable and therefore do not replicate the returns that a hedge fund investor would actually achieve. The returns are likely to differ from the actual returns realized by hedge fund investors due to survivorship bias, backfill bias, and selection bias (Fung and Hsieh 2002). Survivorship bias occurs when the returns of funds that no longer exist are not used in index construction. For the period that we use, HFR tracked deceased funds so survivorship bias may not be as great a concern (Fung and Hsieh 2002). Backfill bias occurs if a hedge fund’s decision to report returns to the HFR database depends on the fund’s past returns and the fund’s past returns are used in the construction of the index. Selection bias is generated by the fact that hedge funds that choose to report to the HFR database are not representative of the universe of all hedge funds.

These biases imply that the average returns of the HFR indices should be interpreted with caution. However, the indices still provide a useful proxy of the risk characteristics of hedge funds in a particular strategy group. In interpreting the results, we focus on correlations (betas) rather than on level differences in returns (alpha). For the monthly analysis we use six HFR indices: Distress/Restructuring, Short Bias, Equity Hedge, Event-Driven, Relative Value, and Convertible Arbitrage. We selected these indices because they have the longest history and because Agarwal and Naik (2004) use them.
3.2.2 Standard Risk Factors

We include a variety of risk factors in four broadly defined areas: equity markets, fixed-income markets, commodity markets, and options. When possible we use the same indices as Agarwal and Naik (2004).

The indices cover US equities (Standard and Poor 500 index and Russell 3000 index), developed-market equities (Morgan Stanley Capital International EAFE index), emerging market equities (MSCI Emerging Markets index), commodities (Goldman Sachs Commodity index), US high yield fixed income (Bank of America Merrill Lynch US High Yield Master II index), US corporate fixed income (Moody’s Corporate BAA Yield index), and the Federal Reserve Bank competitiveness-weighted dollar index. These are the same factors used by Agarwal and Naik (2004) with the exception of the fixed income risk indices (the Lehman High Yield index, the Salomon Brothers government and corporate bond index, and the Salomon Brothers world government bond index). We are not able to obtain data for these throughout the extended sample period.

Additionally, we include the three Fama and French equity market factors: “small-minus-big,” which is the average return on a small cap stock portfolio minus the average return on a large cap portfolio, “high-minus-low,” which is the average return on a value portfolio minus the average return on a growth portfolio, and “up-minus-down,” which is the average return on a portfolio consisting of stocks with high prior-period returns minus the average return on a portfolio consisting of stocks with low prior-period returns.

3.2.3 Risk Factors Based on Options

To test for non-linearities in the exposures of hedge fund returns to the S&P500, we include two types of option-based factors. The first type follows Agarwal and Naik (2004) by trading short-term (ie, 30 to 60 day) at-the-money (ATM) and out-of-the-money (OTM)
put and call options each month. Specifically, the monthly return for the at-the-money put strategy is calculated using the following strategy. On the first trading day of the month, buy the S&P 500 put expiring during the next month with strike price closest to the spot price of the S&P 500. On the first trading day of the following month, sell the put and buy a new put expiring during the following month. Repeating this strategy yields a series of monthly returns for the at-the-money put strategy. Returns for the at-the-money call strategy are calculated similarly. For the out-of-the-money put (call) factor, we employ a similar procedure, but instead choose the option with the strike price that is one tick smaller (larger) than the at-the-money strike. Similar to Agarwal and Naik (2004), for the out-of-the-money options, the average strike-spot ratio upon purchase is 1.01 (or 1/1.01) and it is 1 for the at-the-money options.

One potentially undesirable feature of these OTM option strategies is that they prescribe a relatively constant amount of moneyness (strike to spot ratio), while volatility is certainly time-varying. For example, the risk profile of a 30-day put option that is one percent out of the money is quite different when annual volatility is 30 percent, relative to when volatility is 10 percent. In other words, these strategies are more exposed to time-varying volatility than necessary. To control for this, we construct a second type of option factor, one that employs a trading strategy that buys deeper out-of-the-money puts when forecasted volatility is higher. Specifically, on the first trading day of each month, the strategy purchases the put whose strike price $K$ most closely matches the strike price corresponding to a fixed $K(Z)$:

$$K(Z) = S \cdot \exp\{(r_f(\tau) - q(\tau) + \frac{\sigma^2}{2})\tau + \sigma(\tau)\sqrt{\tau}Z\},$$

where $\sigma$ is defined as forecasted (annualized) volatility, $q$ is dividend yield, $r_f$ is the risk-free rate, $S$ is the index price, and $\tau$ is the time to maturity (in years). Intuitively, a $Z$-
score of negative one corresponds to an option with strike price approximately one standard deviation out-of-the-money, where standard deviation is given by forecasted volatility. We calculate forecasted volatility as 0.8 times the current level of the VIX volatility index.\textsuperscript{11} This fixed-Z strategy exactly replicates that of Jurek and Stafford (2011).

Table 11 presents summary statistics for the monthly risk factors and hedge fund indices. As expected, all option strategies have a strongly negative mean return. This is related to the well-known volatility premium: volatility implied by option prices falls short of realized volatility. The effect is stronger for the out-of-the-money options (implied volatility is higher for out-of-the-money options).

### 3.3 Results of Monthly Risk Factor Analysis

We estimate regressions of the form:

\[
\tilde{r}_t^i = \alpha + \sum_k \gamma_{k,t} \tilde{f}_t^k + \epsilon_{it}
\]

where \(\tilde{r}_t^i\) is the excess return during month \(t\) computed using the values of hedge fund index \(i\), \(\tilde{f}_t^k\) is the excess return during month \(t\) of factor \(k\), and \(\epsilon_{it}\) is an error term uncorrelated with the factors. Excess returns are computed in excess of the risk-free rate as approximated by the return on one-month US treasury bills.

Table 12 presents results using specifications and samples similar to those used by Agarwal and Naik (2004). The main differences are that we use different indices to proxy for fixed income strategies, and that our sample is January 1997 through June 2000 (because we do not have risk factors for the full period of January 1994 through June 2000).\textsuperscript{11} As a robustness check, Agarwal and Naik (2004) employ similar fixed-Z strategies. A major difference between their implementation and this one is that they use realized historical daily volatility (over the last 30 days), where as we use current implied volatility (VIX).
Despite these differences, the main results from Agarwal and Naik (2004) hold. Specifically, restructuring, event-driven, and relative value arbitrage hedge fund strategies have a statistically significant negative exposure to the out-of-the-money put option strategy. And the equity hedge strategy has a statistically significant negative exposure to the at-the-money put option strategy. Short selling strategies, however, do not show a significant exposure to the out-of-the-money call.

Table 13 presents the results with the full sample of January 1997 to March 2011. Despite a fall in standard errors, the out-of-the-money put option strategies are no longer significant in any of the regressions. To investigate why this may have occurred, we examine the returns graphically. Figure 17 plots monthly returns for the three hedge fund strategies against the returns of the S&P 500. Each plot includes the fitted values from a non-parametric regression of the hedge fund returns on the S&P 500 returns. The relationship between the hedge fund returns and S&P 500 returns appears non-linear, especially for the extreme outliers corresponding to the 2008 financial crisis and the unwind of LTCM in 1998. We find it surprising that the returns on the put options are not significant in the regressions given the fact that extending Agarwal and Naik (2004)’s sample through March 2011 adds three “extreme” months in which the S&P 500 lost more than 10%, compared to only one extreme month in the shorter sample. Furthermore, the plots suggest that at least for extreme tail events (LTCM and the financial crisis), hedge fund returns are nonlinear with respect to the return on the S&P 500 with the relationship resembling that of a short put.

One possibility is that even though the hedge fund returns resemble the returns on a short put, the regressors that Agarwal and Naik (2004) selected for each hedge fund strategy better fit the hedge fund returns. Estimating factor regressions for restructuring, event-driven, and relative value arbitrage with only the S&P 500 and put option return as factors yields a negative coefficient on the put option that is significant at the
10% level in all regressions. Another possibility is that the out-of-the-money put option is not far enough out of the money to capture the non-linearity in hedge fund returns. Table 14 presents the same regressions but with a put option with a Z-score of negative one. The option enters significantly at the 5% level in four out of six of the regressions and the sign is negative in five. The coefficient is significant at the 1% or 0.1% level for the three hedge fund strategies that Agarwal and Naik (2004) found were significantly correlated with a short put option (restructuring, event-driven and relative value arbitrage).

3.4 Daily Returns and ETFs

In this section we extend the analysis in two important ways. First, we employ daily instead of monthly return data. Daily data became available relatively recently (in 2003 for the hedge fund returns) and few studies have made use of it. Second, we replace the standard risk factors, many of which are not directly investible, with exchange traded funds, which are liquid instruments designed with the explicit purpose of providing investment exposure to stated risk factors. The fact that ETFs are also relatively new products (especially more exotic ETFs) means that they are also under-utilized in the literature.

This analysis offers a number of advantages and disadvantages relative to our monthly analysis. One obvious advantage is a dramatic increase in the number of observations. Daily hedge fund returns, and daily returns on a large subset of our risk factors, are available starting in 2003: this gives us a total of 2,024 observations for each index, compared to 169 monthly observations for each index in our monthly analysis. Of course, this advantage is partially offset by the fact that the sample only runs from 2003 to 2011, compared to 1996 to 2011 for the monthly data: this means the sample period of our

12Section 3.2.3 details the construction of the put option strategy.
daily analysis misses the important outlier events such as the East Asian/Russian/LTCM events of 1997-1998 and the dot-com bubble events of 2001. Another concern is the accuracy of daily hedge fund returns. Since many hedge funds employ strategies that are at least somewhat illiquid, it is likely difficult to value their day-to-day performance. Moreover, a hedge fund may be tempted to smooth its returns by under-reporting daily losses and gains. Such behavior is difficult or impossible to detect.

Daily ETF returns likewise offer advantages and disadvantages over monthly index returns. Like the hedge fund returns, daily ETF returns trade a shortened time series for more observations. However, this tradeoff is more severe for some ETFs in our sample: for example, data for our high-yield corporate debt ETF (HYG) is only available beginning in 2007. ETFs have the advantage over the traditional risk factors employed by Agarwal and Naik (2004) and others insofar as they are highly liquid and directly investible. The fact that they are investible makes them more relevant for “hedge fund replication” analysis ala Hasanhodzic and Lo (2007).

A final concern about the daily return analysis is its relevance. The majority of hedge funds in our analysis are employing strategies that take more than one day to realize their returns. Moreover, investors are not moving money into and out of hedge funds on a daily basis. So why should we be concerned with daily return correlations? One possibility is that daily return correlations may give us some idea about how returns and factors are correlated over longer periods. Moreover, since our interest is in non-linear risk exposures, observing outliers is crucial for our analysis. By using daily data, we are able to observe more outliers simply because we have more observations.

3.4.1 Hedge Fund Returns

We employ five specific HFRX indices to measure hedge fund returns. Unlike other HFRX indices and the HFRI indices, these five have daily return data. The five indices are con-
structed to replicate the representative returns of five distinct hedge-fund strategy classes: event-driven; relative value; convertible arbitrage; equity hedge; and macro. We employ daily return series for these funds over the period beginning in April 2003 and ending in April 2011, for a total of 2,070 observations for each series. Panel A of Table 15 summarizes our five hedge fund indices.

3.4.2 Risk Factors

An important question is whether the risk exposures (and thus the risk premia) captured by hedge fund strategies can be replicated by liquid securities. Exchange traded funds, or ETFs, are a relatively recent investment product that propose to do precisely this by providing highly liquid and transparent exchange traded investment vehicles that provide exposure to various risk factors. ETF returns are thus a natural set of explanatory variables.

ETF popularity has grown dramatically over the last decade. In 2000, only 92 ETFs were available world-wide, with total assets under management of $74 billion. As of February 2012, 3145 ETFs were available world-wide, with a total of $1.52 trillion assets under management. The breadth of ETF investment strategies has also dramatically increased: while the vast majority of early ETFs were equity products, ETFs are now available that provide exposure to fixed-income products, commodities, currencies, and volatility. The fact that exotic ETFs are so recent means that little academic work has been done comparing ETF and hedge fund risk exposures.

For our risk analysis, we choose a set of ETFs that span a variety of asset classes. For equities, we include four ETFs to assess exposure to large-cap US equities (SPY, the SPDR fund that tracks the S&P500), small-cap US equities (IWM, iShares Russell 2000), other developed market equities (EFA, iShares MSCI EAFE non-US developed market index), and emerging market equities (EEM, iShares MSCI Emerging Market Index).
These four funds are among the top six most liquid ETFs on the market (as measured by average daily volume traded). For fixed income markets, we include three ETFs to assess exposure to US investment grade corporate debt (AGG, iShares Barclay’s Aggregate Investment Grade Bond Index), US high-yield corporate debt (HYG, iShares IBoxx Liquid High Yield Index), and US real estate (VNQ, Vanguard MSCI REIT Index). For commodities, we include two ETFs to measure exposure to gold (GLD, SPDR Gold Trust) and oil (USO, which tracks the price of WTI crude oil). Finally, for currency, we include a single ETF (ERO) to track the performance of the Euro, relative to the dollar. Unfortunately, ETFs that provide exposure to volatility are a very recent innovation. For example, the iPath VXX fund, which takes long positions in short-term VIX futures, is the most popular volatility ETF and was released in January of 2009. Given the relatively short time series of volatility ETF returns, we do not include a volatility ETF.

We follow the literature and our previous analysis by including option-based risk factors. Specifically, we include the returns of two option-trading strategies that take monthly long positions in 30-day at-the-money and out-of-the-money European put options on the S&P500 index. We construct the ATM strategy as follows. Starting on the Monday of the third week of each month, we purchase the put option with strike price closest to the current level of the S&P500, that expires on the third Saturday of the following month (ie, at the close of the market on Friday).13 If no such option is traded on Monday, we look on Tuesday, and so forth. Once the option is identified and purchased, daily returns are constructed for each trading day until the next option is purchased, or the option expires.

13The vast majority of short-dated index options expire at market close on the third Friday of the month. The standard approach in the literature (eg, Agarwal and Naik (2004), or Jurek and Stafford (2011)) is to purchase the option on the 1st of each month (approximately 50 days before expiration) and then sell it on the 1st of the following month (approximately 20 days before expiration). This is because the standard approach is concerned with monthly returns; in contrast, our analysis deals with daily returns, which allows us to employ a more direct option strategy.
To construct the OTM put option strategy, we take a similar approach. The key additional choice variable is the moneyness of the option. Instead of considering options that are a fixed percentage out of the money (for example, one percent out of the money), we follow Jurek and Stafford (2011) by allowing the degree of moneyness to vary with forecasted volatility, where our volatility forecast is 0.8 times the current level of the VIX volatility index.\textsuperscript{14} Our strategy is thus to purchase options that are a fixed number of standard deviations (as measured by forecasted volatility) out of the money; we let this fixed value equal -0.2.\textsuperscript{15} Option selection then follows the same algorithm as in the ATM case, as does the construction of option returns.

The final component of our risk factor set consists of the three Fama-French factors: small minus big (SMB), high minus low (HML), and up minus down (UMD). These are the daily returns of portfolios that are long small-cap and short large cap US equities (SMB), long high book-to-market ratio and short low book-to-market ratio US equities (HML), and long past winners and short past losers (UMD). We obtain the daily returns of these portfolios from Kenneth French’s website.

Panel B of Table 15 summarizes our 15 risk factors, which we divide into three groups. The first group contains three factors: the S&P500 at-the-money put option strategy, the S&P500 out-of-the-money put option strategy, and the S&P500 index itself. These factors are grouped together for two reasons: because data on all three is available throughout the entire HFRX sample period (April 2003 to April 2011), and because these factors will comprise one of our main regression specifications. As is well known, both put strategies deliver negative mean returns over the sample period; this result is directly

\textsuperscript{14}This is described in detail in Section 3.2.3.

\textsuperscript{15}There is a tradeoff when choosing this “Z-score”. Too low of a Z-score (in absolute value) results in a set of option returns that are too closely correlated to the ATM returns. Too high of a Z-score, on the other hand, causes the algorithm to purchase options that are illiquid (options that are closer to the money are more frequently traded) and thus have a large fraction of daily returns equal to zero.
related to the so-called “volatility premium”. The Table also shows that the out-of-the-money strategy has a lower expected return than the at-the-money strategy; this is related to the volatility skew. The second set of risk factors is comprised of four highly liquid ETFs (EEM, IWM, EAF, and AGG) and the three Fama-French factors (SMB, HML, and UMD). Data for these factors begins in September 2003. Of these factors, emerging market equities is the strongest performer over our sampler period, with a mean daily return of 9 basis points. Finally, our last group contains three “exotic” ETFs that provide exposure to junk grade corporate debt (HYG), gold (GLD), and Euro/dollar (ERO). Data for these begins in December of 2007. The top performer in this group is gold, which returned an average of 9 basis points per day over the sample period.

3.5 Results of Daily Risk Factor Analysis

We follow the methods employed in Section 3.3 and perform regressions of the following form:

$$\tilde{r}_i^t = \alpha + \sum_{k \in K(i)} \gamma_k \tilde{f}_i^k + \epsilon_{it}$$

where $\tilde{r}_i^t$ is the excess return during day $t$ computed based on the values of hedge fund index $i$, $K(i)$ is the subset of factors chosen for index $i$, $\tilde{f}_i^k$ is the excess return during day $t$ of factor $k$, and $\epsilon_{it}$ is an error term uncorrelated with the factors. Excess returns are computed in excess of the risk free rate as approximated by the return on one-month US treasury bills. As in our monthly data analysis, we choose a distinct subset of risk factors for each hedge fund index. Unlike our monthly analysis, which replicated the analysis of Agarwal and Naik (2004) in its factor choices, we choose five to seven risk factors per index to test what we feel are intuitive risk exposures for each index. The
specific factor choices and regression results for each index are discussed below.

3.5.1 Equity Hedge

Hedge funds employing equity hedge strategies hold primarily equities and maintain both long and short positions. They vary widely in their net exposures. For example, both 130-30 and 200-200 long-short positions would be included under the equity hedge category. As our regressors we therefore include a variety of equity-based factors: the S&P500 index fund, the out-of-the-money S&P500 put option strategy, the Russell 2000 fund, the non-US developed market equities fund, and the value-growth and small-big Fama-French factors high-minus-low and small-minus-big. Table 17 presents the results of the regression. Interestingly, the coefficient on SPY is not significant, suggesting that the net long-short exposure of the equity hedge index is neutral. However, the coefficient on the OPUT factor is significantly negative, implying that while these positions may be linearly hedged, they may contain second-order (gamma) exposure. For example, many pair-trading strategies are hedged against small market gains and losses, but may incur losses when the market experiences tail losses. For example, a well known trade of Long Term Capital Management involved trading the equities of the dual-listed company Royal Dutch Shell, a trade which should be neutral to small moves in the market, but one which incurred devastating losses in the tail events of 1998. The index also has statistically significant exposures to small-minus-big (positive) and high-minus-low (negative). One way of interpreting this is that hedged equity positions in these funds are net long small cap versus large cap, and net long growth versus value.
3.5.2 Event-driven

Hedge funds employing event-driven strategies hold securities, including both equity and debt, of firms involved in a variety of corporate transactions, such as M&A or financial distress. For the event-driven risk factors, we choose the S&P500 index, the S&P500 put-option strategy, the investment-grade corporate bond fund, the high-yield corporate bond fund, and the Fama-French factors small-minus-big, high-minus-low, and up-minus-down. All six factors have statistically significant coefficients. As with equity hedge, the coefficient on the S&P500 put strategy is significant and negative. This is consistent with event-driven hedge funds taking convergence trades with (short) option-like returns to capitalize on corporate transactions. For example, trading the equities of two firms involved in a proposed merger delivers a small return in the likely event that the merger goes through, but can deliver a large loss if the merger is canceled or renegotiated at an unfavorable share swap ratio.

3.5.3 Relative Value

Relative value hedge funds employ strategies that seek to capitalize on the correction of a pricing discrepancy between related securities; the securities may be equities, fixed-income, or other. For our risk factors, we choose the S&P500 index fund, its put option strategy, and the three Fama-French factors high-minus-low, small-minus-big, and up-minus-down. The only factor with a statistically significant coefficient is the put factor factor, whose coefficient is negative. Like equity hedge and event-driven, relative value hedge funds may employ strategies that are delta-neutral to the market, but that experience negative gamma (for losses).
3.5.4 Convertible Arbitrage

Convertible arbitrage hedge funds hold securities whose relative prices experience a discrepancy from some inherent relationship, such as convertible bonds and the associated equity. For our risk factors we choose the S&P500 index, its put strategy, the investment-grade corporate bond fund, the high-yield corporate bond fund, and the Fama-French factors high-minus-low, small-minus-big, and up-minus-down. Interestingly, the coefficient on the S&P500 index is significant and negative, while the coefficients on the investment-grade and high-yield bond funds are significant and positive. This is consistent with a trade that is long convertible debt and short the associated stock.

3.5.5 Macro

Macro strategies typically seek to capitalize on future movements in various macro variables by holding equities, fixed-income securities, currencies, or commodities, and their respective futures and derivatives. For risk factors we choose the S&P500 index, its put strategy, the emerging market equity fund, the Fama-French momentum factor, the gold fund, and the Euro/dollar fund. The coefficient on the emerging markets equity fund is significant and positive, while the coefficient on the S&P500 is significant and negative. The macro index also experiences positive loading on the momentum (long past winners, short past losers) trade. Unlike the equity hedge, event-driven, and relative value indices, macro does not have a significantly negative correlation with the S&P500 put strategy.

3.6 Conclusion

We extend existing analyses of hedge fund risk exposures in two ways: by employing monthly data through 2011 (including the financial crisis), and by employing daily data. In contrast to Agarwal and Naik (2004), we do not find that hedge fund returns are sig-
nificantly correlated with at-the-money options; however, we do find exposure to deeper out-of-the-money puts, similar to the findings of Jurek and Stafford (2011). One final contribution is our use of exchange-traded funds as risk factors, relatively new investment vehicles that are under-utilized in the replication literature. A promising line of future research would be a further investigation of whether hedge fund returns can be replicated by static or dynamic ETF trading strategies.
Panel A: Hedge Fund Indexes

<table>
<thead>
<tr>
<th>Hedge fund strategy</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restructuring</td>
<td>0.77</td>
<td>1.94</td>
<td>0.94</td>
<td>-1.50</td>
<td>8.37</td>
<td>-8.50</td>
<td>5.55</td>
</tr>
<tr>
<td>Event driven</td>
<td>0.79</td>
<td>2.09</td>
<td>1.11</td>
<td>-1.30</td>
<td>7.04</td>
<td>-8.90</td>
<td>5.13</td>
</tr>
<tr>
<td>Relative Value</td>
<td>0.68</td>
<td>1.35</td>
<td>0.80</td>
<td>-3.00</td>
<td>18.20</td>
<td>-8.90</td>
<td>3.93</td>
</tr>
<tr>
<td>Convertible Arbitrage</td>
<td>0.70</td>
<td>2.28</td>
<td>0.88</td>
<td>-2.90</td>
<td>25.50</td>
<td>-9.00</td>
<td>9.74</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>0.83</td>
<td>2.85</td>
<td>0.92</td>
<td>-0.10</td>
<td>4.93</td>
<td>-9.00</td>
<td>10.90</td>
</tr>
<tr>
<td>Equity Short Biased</td>
<td>0.20</td>
<td>5.59</td>
<td>-0.20</td>
<td>0.28</td>
<td>5.96</td>
<td>-21.00</td>
<td>22.80</td>
</tr>
</tbody>
</table>

Panel B: Risk Factors

<table>
<thead>
<tr>
<th>Risk factor name</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSCI Non-US/Canada Developed Countries</td>
<td>0.32</td>
<td>5.14</td>
<td>0.84</td>
<td>-0.70</td>
<td>4.32</td>
<td>-20.60</td>
<td>12.50</td>
</tr>
<tr>
<td>MSCI Emerging Markets</td>
<td>0.59</td>
<td>7.50</td>
<td>0.96</td>
<td>-0.73</td>
<td>4.54</td>
<td>-28.20</td>
<td>17.40</td>
</tr>
<tr>
<td>FRB Competitiveness-weighted Dollar</td>
<td>-0.01</td>
<td>1.32</td>
<td>-0.11</td>
<td>0.89</td>
<td>6.29</td>
<td>-3.26</td>
<td>6.64</td>
</tr>
<tr>
<td>Small-minus-big (SMB)</td>
<td>0.27</td>
<td>3.86</td>
<td>-0.01</td>
<td>0.83</td>
<td>10.40</td>
<td>-16.60</td>
<td>22.10</td>
</tr>
<tr>
<td>High-minus-low (HML)</td>
<td>0.36</td>
<td>3.70</td>
<td>0.31</td>
<td>-0.02</td>
<td>5.20</td>
<td>-12.90</td>
<td>13.90</td>
</tr>
<tr>
<td>Risk free</td>
<td>0.25</td>
<td>0.17</td>
<td>0.27</td>
<td>-0.06</td>
<td>1.51</td>
<td>0.00</td>
<td>0.56</td>
</tr>
<tr>
<td>Up-minus-down (UMD)</td>
<td>0.57</td>
<td>6.12</td>
<td>0.98</td>
<td>-1.49</td>
<td>10.80</td>
<td>-34.80</td>
<td>18.40</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>0.43</td>
<td>4.79</td>
<td>1.00</td>
<td>-0.66</td>
<td>3.65</td>
<td>-16.90</td>
<td>9.67</td>
</tr>
<tr>
<td>Goldman Sachs Commodity</td>
<td>0.75</td>
<td>6.93</td>
<td>1.42</td>
<td>-0.36</td>
<td>4.14</td>
<td>-27.80</td>
<td>21.10</td>
</tr>
<tr>
<td>Russell 3000</td>
<td>0.60</td>
<td>4.88</td>
<td>1.35</td>
<td>-0.73</td>
<td>3.83</td>
<td>-17.70</td>
<td>10.50</td>
</tr>
<tr>
<td>Moody’s Corporate BAA</td>
<td>-0.11</td>
<td>3.47</td>
<td>-0.30</td>
<td>1.54</td>
<td>12.00</td>
<td>-10.60</td>
<td>21.50</td>
</tr>
<tr>
<td>Merrill High Yield</td>
<td>0.60</td>
<td>2.90</td>
<td>0.95</td>
<td>-1.23</td>
<td>11.20</td>
<td>-16.30</td>
<td>11.50</td>
</tr>
<tr>
<td>S&amp;P 500 at-the-money call</td>
<td>24.90</td>
<td>83.20</td>
<td>59.30</td>
<td>1.68</td>
<td>5.59</td>
<td>-97.20</td>
<td>320.00</td>
</tr>
<tr>
<td>S&amp;P 500 out-of-the-money put</td>
<td>29.70</td>
<td>84.40</td>
<td>65.60</td>
<td>1.87</td>
<td>6.29</td>
<td>-98.00</td>
<td>331.00</td>
</tr>
<tr>
<td>S&amp;P 500 very out-of-the-money put (Z = -1)</td>
<td>-33.70</td>
<td>89.60</td>
<td>72.60</td>
<td>2.33</td>
<td>9.24</td>
<td>-100.00</td>
<td>438.00</td>
</tr>
<tr>
<td>S&amp;P 500 at-of-the-money put</td>
<td>-10.20</td>
<td>68.50</td>
<td>21.50</td>
<td>0.52</td>
<td>2.21</td>
<td>-99.80</td>
<td>167.00</td>
</tr>
</tbody>
</table>

Table 11: Monthly Summary Statistics. Summary statistics of monthly returns of Hedge Fund Research Indices and risk factors. The sample is from January 1997 through March 2011, excluding July 1997 and October to December 2001 due to missing data. See the text for a description of risk factors and hedge fund indices.
### Table 12: Risk Factor Analysis Partial Sample: Monthly Returns, January 1997 to June 2000. Each column shows results from a separate regression of the excess returns (above the risk-free rate) of an HFR index on the excess returns of the risk factors. Standard errors in parenthesis. ∗ ∗ ∗ * p < .001, ∗ ∗ p < .01, ∗ p < .05. The hedge fund indices are Distressed/Restructuring (REST), Event-Driven (ED), Relative Value (RVAL), Convertible Arbitrage (CA), Equity Hedge (EH), and Short Bias (SHORT). The risk factors (rows) are the out-of-the-money put option described in the paper (SPPo), the at-the-money put option (SPPa), the out-of-the-money call option (SPCo), Russell 3000 index (RUS), lagged Russell 3000 index (LRUS), MSCI non-US/Canada Developed Country Index (MSCI EAFE), MSCI Emerging Markets Index (MEM), Federal Reserve Bank competitiveness-weighted Dollar Index (FRBI), Fama-French size and value factors (SMB, HML), Momentum factor (UMD), S&P 500 Index (SP500), Goldman Sachs commodity index (GSCI), Bank of America Merrill Lynch US High Yield Master II index (MHY), and Moody’s Corporate BAA Yield index (MOOD). The sample does not include July 1997 due to missing options data.
<table>
<thead>
<tr>
<th>Restructuring</th>
<th>Event driven</th>
<th>Relative value arbitrage</th>
<th>Convertible arbitrage</th>
<th>Equity hedge</th>
<th>Short selling</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.298***</td>
<td>-0.175***</td>
<td>0.00319***</td>
<td>0.157**</td>
<td>-0.390***</td>
</tr>
<tr>
<td>(0.0457)</td>
<td>(0.0350)</td>
<td>(0.00893)</td>
<td>(0.0486)</td>
<td>(0.0210)</td>
<td>(0.0622)</td>
</tr>
<tr>
<td>SPPo</td>
<td>-0.00262</td>
<td>-0.00152</td>
<td>-0.00341*</td>
<td>-0.00987</td>
<td>RUS 0.394***</td>
</tr>
<tr>
<td>(0.00148)</td>
<td>(0.00172)</td>
<td>(0.00135)</td>
<td>(0.00204)</td>
<td>(0.0209)</td>
<td>(0.00432)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.0811**</td>
<td>0.156***</td>
<td>-0.00719</td>
<td>0.0955***</td>
<td>SMB 0.201***</td>
</tr>
<tr>
<td>(0.0265)</td>
<td>(0.0232)</td>
<td>(0.0142)</td>
<td>(0.0272)</td>
<td>(0.0269)</td>
<td>(0.0616)</td>
</tr>
<tr>
<td>HML</td>
<td>0.0219</td>
<td>0.0494*</td>
<td>0.0439</td>
<td>-0.00706</td>
<td>HML -1.05***</td>
</tr>
<tr>
<td>(0.0280)</td>
<td>(0.0239)</td>
<td>(0.0225)</td>
<td>(0.0349)</td>
<td>(0.0282)</td>
<td>(0.0465)</td>
</tr>
<tr>
<td>LRUS</td>
<td>0.108***</td>
<td>0.179***</td>
<td>0.0344</td>
<td>0.123***</td>
<td>GSCI 0.0832***</td>
</tr>
<tr>
<td>(0.0209)</td>
<td>(0.0239)</td>
<td>(0.0219)</td>
<td>(0.0234)</td>
<td>(0.0145)</td>
<td>(0.0145)</td>
</tr>
<tr>
<td>MHY</td>
<td>0.193***</td>
<td>0.0973***</td>
<td>0.117***</td>
<td>MOOD -0.249***</td>
<td></td>
</tr>
<tr>
<td>(0.0453)</td>
<td>(0.0172)</td>
<td>(0.0224)</td>
<td>(0.0384)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRBI</td>
<td>-0.0656</td>
<td>MEM 0.0760***</td>
<td>MSCIEAFE</td>
<td>0.0973***</td>
<td></td>
</tr>
<tr>
<td>(0.0833)</td>
<td></td>
<td>(0.0186)</td>
<td>(0.0273)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEM</td>
<td>0.0760***</td>
<td>Obs 169</td>
<td>Obs 169</td>
<td>Obs 169</td>
<td>Obs 169</td>
</tr>
<tr>
<td>(0.0186)</td>
<td></td>
<td>Adj-R-Sq 0.749</td>
<td>Adj-R-Sq 0.423</td>
<td>Adj-R-Sq 0.479</td>
<td>Adj-R-Sq 0.809</td>
</tr>
<tr>
<td>Adj-R-Sq</td>
<td>0.650</td>
<td>Obs 169</td>
<td>Obs 169</td>
<td>Obs 169</td>
<td>Obs 169</td>
</tr>
</tbody>
</table>


* * *p < .001, * * p < .01, * p < .05. The hedge fund indices are Distressed/Restructuring (REST), Event-Driven (ED), Relative Value (RVAL), Convertible Arbitrage (CA), Equity Hedge (EH), and Short Bias (SHORT). The risk factors (rows) are the out-of-the-money put option described in the paper (SPPo), the at-the-money put option (SPPa), the out-of-the-money call option (SPCo), Russell 3000 index (RUS), lagged Russell 3000 index (LRUS), MSCI non-US/Canada Developed Country Index (MSCI EAFE), MSCI Emerging Markets Index (MEM), Federal Reserve Bank competitiveness-weighted Dollar Index (FRBI), Fama-French size and value factors (SMB, HML), Momentum factor (UMD), S&P 500 Index (SP500), Goldman Sachs commodity index (GSCI), Bank of America Merrill Lynch US High Yield Master II index (MHY), and Moody’s Corporate BAA Yield index (MOOD). The sample does not include July 1997 or October to December 2001 due to missing options data.
Table 14: Risk Factor Analysis Full Sample with Very Out-of-the-Money Put: Monthly Returns, January 1997 to March 2011. Each column shows results from a separate regression of the excess returns (above the risk-free rate) of an HFR index on the excess returns of the risk factors. Standard errors in parenthesis. **p < .001, *p < .01, p < .05. The hedge fund indices are Distressed/Restructuring (REST), Event-Driven (ED), Relative Value (RVAL), Convertible Arbitrage (CA), Equity Hedge (EH), and Short Bias (SHORT). The risk factors (rows) are the very out-of-the-money put option strategy (Z = 1) described in the paper (SPPvo), Russell 3000 index (RUS), lagged Russell 3000 index (LRUS), MSCI non-US/Canada Developed Country Index (MSCI EAFE), MSCI Emerging Markets Index (MEM), Federal Reserve Bank competitiveness-weighted Dollar Index (FRBI), Fama-French size and value factors (SMB, HML), Momentum factor (UMD), S&P 500 Index (SP500), Goldman Sachs commodity index (GSCI), Bank of America Merrill Lynch US High Yield Master II index (MHY), and Moody’s Corporate BAA Yield index (MOOD). The sample does not include July 1997 or October to December 2001 due to missing options data.
### Panel A: Hedge Fund Indexes

<table>
<thead>
<tr>
<th>Hedge fund strategy</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event driven</td>
<td>0.02</td>
<td>0.31</td>
<td>0.04</td>
<td>-1.45</td>
<td>14.20</td>
<td>-3.12</td>
<td>1.88</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Relative Value</td>
<td>0.01</td>
<td>0.30</td>
<td>0.03</td>
<td>-2.25</td>
<td>32.50</td>
<td>-3.69</td>
<td>2.57</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Convertible Arbitrage</td>
<td>-0.02</td>
<td>0.45</td>
<td>0.01</td>
<td>-4.56</td>
<td>51.00</td>
<td>-6.43</td>
<td>3.28</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>0.01</td>
<td>0.42</td>
<td>0.05</td>
<td>-0.82</td>
<td>7.66</td>
<td>-3.02</td>
<td>2.56</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Macro</td>
<td>0.01</td>
<td>0.45</td>
<td>0.03</td>
<td>-0.97</td>
<td>9.18</td>
<td>-3.65</td>
<td>2.01</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
</tbody>
</table>

### Panel B: Risk Factors

<table>
<thead>
<tr>
<th>Risk factor name</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P at-the-money put</td>
<td>-5.54</td>
<td>43.40</td>
<td>-6.75</td>
<td>1.83</td>
<td>12.40</td>
<td>-100.00</td>
<td>400.00</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>S&amp;P out-of-the-money put</td>
<td>-6.03</td>
<td>61.90</td>
<td>-13.00</td>
<td>7.22</td>
<td>106.00</td>
<td>-91.70</td>
<td>1236.00</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>0.03</td>
<td>1.28</td>
<td>0.09</td>
<td>-0.32</td>
<td>14.20</td>
<td>-9.85</td>
<td>11.70</td>
<td>4/1/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Emerging Market Equities ETF</td>
<td>0.09</td>
<td>2.22</td>
<td>0.16</td>
<td>0.28</td>
<td>14.10</td>
<td>-16.20</td>
<td>20.90</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Russell 2000 ETF</td>
<td>0.04</td>
<td>1.68</td>
<td>0.13</td>
<td>-0.26</td>
<td>7.63</td>
<td>-11.20</td>
<td>8.51</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>MSCI Non-US/Canada Developed Country ETF</td>
<td>0.04</td>
<td>1.56</td>
<td>0.10</td>
<td>-0.11</td>
<td>12.60</td>
<td>-11.20</td>
<td>13.80</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>US Investment Grade Bond ETF</td>
<td>0.02</td>
<td>0.36</td>
<td>0.01</td>
<td>-3.47</td>
<td>77.90</td>
<td>-6.83</td>
<td>2.49</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Small-minus-big (SMB)</td>
<td>0.01</td>
<td>0.58</td>
<td>0.02</td>
<td>-0.04</td>
<td>7.67</td>
<td>-3.80</td>
<td>4.31</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>High-minus-low (HML)</td>
<td>0.02</td>
<td>0.62</td>
<td>0.01</td>
<td>0.29</td>
<td>11.00</td>
<td>-3.32</td>
<td>4.03</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Up-minus-down (UMD)</td>
<td>0.00</td>
<td>1.10</td>
<td>0.06</td>
<td>-0.88</td>
<td>13.30</td>
<td>-8.29</td>
<td>7.10</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Risk free</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.56</td>
<td>1.87</td>
<td>0.00</td>
<td>0.02</td>
<td>9/30/2003</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>High Yield US Corporate ETF</td>
<td>0.02</td>
<td>1.12</td>
<td>0.06</td>
<td>-0.53</td>
<td>15.50</td>
<td>-8.09</td>
<td>7.88</td>
<td>12/28/2007</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Long Gold ETF</td>
<td>0.00</td>
<td>1.48</td>
<td>0.14</td>
<td>0.22</td>
<td>9.05</td>
<td>-7.43</td>
<td>11.30</td>
<td>12/28/2007</td>
<td>4/29/2011</td>
</tr>
<tr>
<td>Euro Dollar Carry ETF</td>
<td>0.01</td>
<td>0.98</td>
<td>0.00</td>
<td>0.53</td>
<td>23.50</td>
<td>-8.91</td>
<td>9.51</td>
<td>12/28/2007</td>
<td>4/29/2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Event driven (1)</th>
<th>Relative value arbitrage (2)</th>
<th>Convertible arbitrage (3)</th>
<th>Equity hedge (4)</th>
<th>Macro (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.00000803</td>
<td>-0.0000562</td>
<td>-0.000245*</td>
<td>-0.000102</td>
<td>-0.00000712</td>
</tr>
<tr>
<td></td>
<td>(0.0000558)</td>
<td>(0.0000645)</td>
<td>(0.000101)</td>
<td>(0.0000671)</td>
<td>(0.000101)</td>
</tr>
<tr>
<td>SP500</td>
<td>-0.000801***</td>
<td>-0.000901***</td>
<td>-0.000217</td>
<td>-0.00104***</td>
<td>-0.00102***</td>
</tr>
<tr>
<td></td>
<td>(0.000104)</td>
<td>(0.000121)</td>
<td>(0.000189)</td>
<td>(0.000125)</td>
<td>(0.000188)</td>
</tr>
<tr>
<td>SP500</td>
<td>0.118***</td>
<td>0.0167**</td>
<td>-0.0278**</td>
<td>0.205***</td>
<td>-0.0228*</td>
</tr>
<tr>
<td></td>
<td>(0.00506)</td>
<td>(0.00585)</td>
<td>(0.00917)</td>
<td>(0.00608)</td>
<td>(0.00912)</td>
</tr>
<tr>
<td>Obs</td>
<td>2.024</td>
<td>2.024</td>
<td>2.024</td>
<td>2.024</td>
<td>2.024</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.346</td>
<td>0.054</td>
<td>0.005</td>
<td>0.500</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Table 16: Risk Factor Analysis with Out-of-the-Money Put: Daily Returns, April 1 2003 to April 29 2011. Each column shows results from a separate regression of the excess returns (above the risk-free rate) of an HFR index on the excess returns of the risk factors. Standard errors in parenthesis. ** * *p < .001, ** * p < .01, * p < .05. The hedge fund indices are Event-Driven (ED), Relative Value (RVAL), Convertible Arbitrage (CA), Equity Hedge (EH), and Macro (MACRO). The risk factors (rows) are the out-of-the-money put option strategy described in the paper (SPPo) and the S&P 500 ETF (SP500).
Table 17: Risk Factor Analysis with Out-of-the-Money Put and Selected Risk Factors: Daily Returns. Each column shows results from a separate regression of the excess returns (above the risk-free rate) of an HFR index on the excess returns of the risk factors. Standard errors in parenthesis. ∗∗∗ \( p < .001 \), ∗∗ \( p < .01 \), ∗ \( p < .05 \). The hedge fund indices are Event-Driven (ED), Relative Value (RVAL), Convertible Arbitrage (CA), Equity Hedge (EH), and Macro (MACRO). The risk factors (rows) are the out-of-the-money put option strategy described in the paper (SPPo), the S&P 500 ETF (SP500), Emerging Market Equities ETF (EEM), Russell 2000 ETF (IWM), MSCI Non-US/Canada Developed Country ETF (EAFE), US Investment Grade Bond ETF (AGG), High Yield US Corporate ETF (HYG), Long Gold ETF (GLD), Euro Dollar Carry ETF (ERO), Fama-French size and value factors (SMB, HML), Momentum factor (UMD).
Figure 16: Cumulative Monthly Hedge Fund Index Returns. This figure plots the cumulative monthly returns of seven HFR indices and the S&P500, between 1997 and 2011. Five out of the seven indices experience substantial losses during the financial crisis; the exceptions are Short-bias and Macro.
Figure 17: Monthly HFRI Returns vs S&P500 Returns. This figure plots the monthly returns of the Restructuring, Event-Driven, and Relative Value HFRI indices against the monthly returns of the S&P500 index from 1997 to 2011. August 1998 (LTCM) and October 2008 (Lehman) are significant outliers.
Figure 18: Daily HFRX Returns vs Daily S&P500 Returns: Equity Hedge, Event-Driven, and Relative Value. Panel (a) plots daily returns of the HFRX Equity Hedge index against daily S&P500 returns. Panel (b) plots daily returns of the HFRX Event-Driven index against daily S&P500 returns. Panel (c) plots daily returns of the HFRX Relative Value index against daily S&P500 returns. The sample includes 2,024 trading days between April 2003 and April 2011.
Figure 19: Daily HFRX Returns vs Daily S&P500 Returns: Convertible Arbitrage and Macro. Panel (a) plots daily returns of the HFRX Convertible Arbitrage index against daily S&P500 returns. Panel (b) plots daily returns of the HFRX Macro index against daily S&P500 returns. The sample includes 2,024 trading days between April 2003 and April 2011.
4 Bibliography


72–85.


Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments
37.

Longstaff, F. A. and Schwartz, E. S. (1995). A simple approach to valuing risky fixed and


