OFR WORKING PAPER
OFFICE OF FINANCIAL RESEARCH

15-11 | May 28, 2015
Revised Dec. 14, 2016

Systemwide Commonalities in Market Liquidity

Mark D. Flood
Office of Financial Research
mark.flood@ofr.treasury.gov

John C. Liechty
Office of Financial Research
john.liechty@ofr.treasury.gov

Thomas Piontek
Office of Financial Research
thomas.piontek@ofr.treasury.gov

Systemwide Commonalities in Market Liquidity

Mark D. Flood † John C. Liechty ‡ Thomas Piontek †

December 14, 2016

† – Office of Financial Research, Washington, DC
‡ – Pennsylvania State University, State College, PA, and Office of Financial Research, Washington, DC

The authors thank Tobias Adrian, Rick Bookstaber, Jeremy Chiu, Greg Feldberg, Pete Kyle, Trish Mosser, Rob Engle, Tuomas Peltonen, Shrihari Santosh, Yves Schüler, and conference and workshop participants at Goethe University, Massachusetts Institute of Technology, European Systemic Risk Board, Federal Reserve Bank of New York, Open Financial Data Group, International Conference of the Financial Engineering and Banking Society (FEBS 2015), International Monetary Fund, International Symposium on Forecasting at the University of California at Riverside (IFS 2015), Office of Financial Research (OFR), the OFR–Federal Reserve Bank of Cleveland 2014 Financial Stability Conference, the New York University finance seminar, the Federal Housing Finance Agency, the Institute for Pure and Applied Mathematics at the University of California at Los Angeles, the 9th International Conference on Computational and Financial Econometrics (CFE 2015) at the University of London, the Cambridge University Centre for Risk Studies, the 2016 Kyle Conference at the University of Maryland, and the 2016 Workshop on Algorithms for Modern Massive Data Sets (MMDS) at the University of California at Berkeley for numerous helpful comments and suggestions. Any remaining errors are the responsibility of the authors alone.

Systemwide Commonalities in Market Liquidity

Abstract

We explore statistical commonalities among granular measures of market liquidity with the goal of illuminating systemwide patterns in aggregate liquidity. We calculate daily invariant price impacts described by Kyle and Obizhaeva [2016] to assemble a granular panel of liquidity measures for equity, corporate bond, and futures markets. We estimate Bayesian models of hidden Markov chains and use Markov chain Monte Carlo analysis to measure the latent structure governing liquidity at the systemwide level. Three latent liquidity regimes — high, medium, and low price-impact — are adequate to describe each of the markets. Focusing on the equities subpanel, we test whether a collection of systemwide market summaries can recover the estimated liquidity dynamics. This version of the model allows an economically meaningful attribution of the latent liquidity states and yields meaningful predictions of liquidity disruptions as far as 15 trading days in advance of the 2008 financial crisis.

Contents

1 Introduction

We present a new approach to the study of commonalities in market liquidity. One goal is to identify broad patterns in granular, daily data that might support a program for monitoring systemwide liquidity. We build up from granular measures of liquidity across a broad range of individual markets, using the recent invariant price-impact measures of Kyle and Obizhaeva [2016] to estimate a daily panel of overall liquidity conditions. Specifically, our initial implementation considers volatility index futures, oil futures, sector portfolios for the Center for Research in Securities Prices (CRSP) universe of U.S. equities, and portfolios based on bond ratings for the Transaction Reporting and Compliance Engine (TRACE) universe of corporate bonds over the decade 2004-14.

The market-invariant approach to price impact carefully normalizes for local volume and volatility conditions to measure liquidity (price impact) in a way that is comparable — i.e., has an invariant measurement scale — across markets and over time. Comparability is crucial for aggregating local liquidity conditions to support systemwide analysis. We calculate a daily measure of liquidity for each security as the price impact of a trade representing one percent of the average daily volume in that market over the preceding month. For futures contracts, we measure liquidity separately for each distinct futures maturity for which data are available, starting with the front-month contract. For stocks and bonds, we aggregate first into portfolios, and then calculate the one-percent price impact within each portfolio. Because liquidity is estimated separately within each market, but measured on an invariant scale, the framework should be readily extensible to broader (or narrower) data panels.

Based on this panel of daily liquidity measurements, we estimate Bayesian hidden Markov chain (HMC) models to capture the latent structure of each individual series, and then apply Markov chain Monte Carlo (MCMC) analysis to assess the latent structure governing liquidity at the systemwide level. The HMC approach posits that the dynamics of each daily price impact measure (28 in our sample) are determined by an underlying variable that alternates among several liquidity states to drive sudden changes in the observed levels of price impact. The underlying states are latent — i.e., not directly observable — but inferred from the dynamics of daily price-impact measurements. In the initial analysis, we estimate each price impact series independently; that is, we assume no coordination between the dynamics of the latent liquidity states across markets. Nonetheless, we find surprising consistency in the dynamics of market liquidity across all of these markets. From the perspective of a policy maker who seeks to identify, or even predict, turbulent episodes in the financial system, we find that three liquidity regimes are adequate to describe each market: high, intermediate, and low. Moreover, we find that the low liquidity regime afflicts all markets roughly simultaneously during the financial crisis of 2008.

We build on earlier studies of liquidity commonalities by linking the estimated latent liquidity states from multiple markets together in a multinomial probit model driven by a daily panel of system-level summary series. This framework permits an assessment of summary series as potential advance indicators of systemic illiquidity. The model reveals that a number of summary series, including the Dow Jones U.S. Real Estate Industry Group Index, Treasury-Eurodollar (TED) spread, VIX®, and the S&P 500 price-to-book (P/B) ratio are statistically significant in explaining liquidity in the equity markets. These summary series are observed daily and unsurprisingly exhibit a high level of temporal correlation. As a result, adding lagged values generally does not improve the performance of the probit model. On the other hand, the same high level of temporal correlation allows a version of the model based on lagged summary series to predict sudden future shifts in the liquidity states — as much as 15 days in advance of the liquidity crisis in September 2008. Although a detailed exploration of the predictive power of these methods is a point of future research, these preliminary results suggest the method might support market monitoring and early warning systems for illiquidity episodes.

The paper is structured as follows. The remainder of Section 1 discusses the general challenges of liquidity measurement, and provides a rationale for our approach. Section 2 describes the models and sampling strategies for the MCMC analysis for both the univariate models (i.e., one market at a time) and the hierarchical model (multiple markets at a time). Section 3 describes the data and specific formulas for measuring price impact. Section 4 reports the findings from aggregating both the market specific analyses and the multiple-market analysis. Section 5 discusses some alternative modeling approaches (e.g., vector autoregressive models) and their limitations with respect to policy applications, and also touches on how these might fit into a broader modeling approach that focuses on prediction of liquidity dynamics; this discussion ends with a demonstration of the predictive power of the probit model during the crisis of 2008. We conclude with a discussion of potential future work, including ways that limitations of the current approach might be overcome to improve our ability to forecast liquidity.

1.1 Overview of the Literature on Market Liquidity

Studies of market liquidity are typically grounded in practical considerations about market quality and frequently exploit local microstructural characteristics to craft market-specific liquidity measures.[1] This tendency towards customization of liquidity metrics tends to support the economic interpretation of results, while making full use of the available data. Customized measurement is a hindrance for a systemwide analysis, however, because data availability and the microstructural interpretation of the estimated measures can vary considerably from one market to the next. As a point of comparison, we consider in Section 2.1 the dynamics of four more traditional measures of liquidity: the implied bid-ask spread

[1] The literature on market liquidity is very extensive, with many recent overviews; e.g., Foucault et al. [2013], Holden et al. [2014], Anderson et al. [2015], Gabrielsen et al. [2011], Vayanos and Wang [2012], Hasbrouck [2007], and Stoll [2000]. Amihud et al. [2013] compile a set of classic papers.

of Roll [1984]; Kyle's-λ (see Kyle [1985]); the volume-scaled absolute return measure of Amihud [2002]; and log volatility. While these alternative measures indeed capture certain patterns in liquidity dynamics, they also exhibit clear statistical anomalies and occasional structural changes related to microstructural innovations. Among the market liquidity measures currently available in the literature, the market-invariant measure most closely meets the requirement of comparability over time and across markets.

Our work builds on earlier studies that look for aggregate liquidity patterns. Chordia et al. [2000] was the first in a series of papers to search for commonalities in liquidity in the cross-section of equity markets. They perform time-series regressions of individual equities' liquidity (measured as market depth and bid-ask spreads) on cross-sectional average measures of liquidity. The data are noisy — R^2s are low — but there is strong evidence of contemporaneous correlation between individual stocks and the aggregate. Karolyi et al. [2012] extend the analysis to an international comparison of thousands of stocks in 40 countries. Again, commonalities in liquidity exist, and unsurprisingly differ significantly across countries and over time.[2] Recent research by Chen et al. [2012] combines price information from financial markets with quarterly quantity information from the Federal Reserve's Flow of Funds data in an effort to distinguish the differential impact of shifts in liquidity demand versus liquidity supply. They distinguish between core and noncore liquidity, where the noncore category includes financial firms' liabilities held by other financial institutions.

We follow Brunnermeier and Pedersen [2009] in distinguishing "market" liquidity for financial assets, such as corporate stocks and bonds, from "funding" liquidity for financial intermediaries' short-term liabilities. From a systemic perspective, funding liquidity in the

[2] Karolyi et al. [2012] define commonality by the R^2 of each stock's daily price-impact measure, per Amihud [2002], on the average price impact for all other stocks in the country. Individual stock commonality measures are averaged to get a country-level commonality index. Karolyi et al. [2012, p. 99] attribute the time-series variation to both supply- and demand-side proxies in funding markets via regression analysis, noting that "demand-side explanations are more reliably significant."

wholesale markets for institutional liabilities is the most immediate concern. However, there are vastly more individual markets for financial assets than for intermediaries' liabilities, and it is an empirical question whether additional information exists in this much larger panel of asset markets that might help to explain liquidity conditions in the funding markets. Previous studies of commonalities in liquidity, cited above, find that there are indeed significant patterns in the detailed data. We expand on this work by analyzing a range of asset classes, including equities, bonds, and financial futures. Most prior studies of commonality have focused on equities markets alone.[3] We also extend the commonalities approach with a novel methodology for connecting aggregate liquidity patterns to a panel of systemwide market summary time series.

Financial institutions aggregate and reallocate liquidity. Systemic imbalances in liquidity therefore tend to appear in wholesale funding markets, where they are a commonplace feature of financial crises. Allen and Gale [2009] distinguish the role in crises of liquidity fundamentals (e.g., subprime mortgage valuations), which operate primarily in asset markets, and panics (e.g., bank runs), which primarily affect funding markets. Both forces might be present in any particular episode, and measuring their influence is an important empirical task. Ideally, measures of liquidity to support financial stability monitoring would be timely — that is, available at high frequency to track developments in near real time. Measures would also be forward-looking — possessing some forecasting power to serve as an early warning signal. These goals are often defeated in practice by certain fundamental challenges. In particular, liquidity exhibits three interrelated characteristics that are obstacles to measurement: latency, nonlinearity, and endogeneity. Each of these challenges has ramifications for both funding liquidity and market liquidity.

Latency means that much liquidity behavior is ex-ante unobserved. We typically most

[3] An exception is the recent working paper by Marra [2013] which pairs individual equities with their matching credit default swaps (CDS); her emphasis, however, is on firm-level interactions between the securities rather than systemwide liquidity.

wish to know how deep or resilient the market will be in the presence of unusual order flow. The most interesting liquidity events are therefore also the rarest. Many trading mechanisms encourage the participation of liquidity providers by restricting information availability, such as closed limit-order books and hidden (or "iceberg") orders; e.g., [Parlour and Seppi, 2008, Section 2.6], Bessembinder et al. [2009]. More generally, markets can utilize anonymous brokerage to conceal trader identities, and/or limited-access upstairs trading venues for large trades; e.g., Degryse et al. [2014], Nimalendran and Ray [2014], Zhu [2013], and Foley et al. [2013]. The inability of researchers to measure these aspects of liquidity directly makes these forces a natural subject for empirical modeling. If key drivers, such as dealer intentions or customer order flow, are not immediately observable, latent structure might be recovered through statistical inference.[4] We address the challenge of latency by adopting a MCMC approach to estimate the latent regime structure governing the observed price impact series. Although the markets' liquidity behavior is indeed largely latent, these hidden patterns may reveal themselves in a broad cross-section of markets observed at relatively high frequency (daily). Our results are indeed able to identify three meaningful latent liquidity states (high, medium, and low price-impact) that seem to govern the observed liquidity behavior.

Nonlinearity in the response of liquidity to significant market changes hampers our ability to extrapolate from small-scale, localized effects to the larger, out-of-sample effects that are often of greatest concern. Numerous studies document the empirical regularity that price response to order flow tends to be concave function of the transaction size. Intuitively, order flow can move the price significantly before additional liquidity providers arrive to dampen the effect. Much of the literature identifies a square-root rule that posits price impact to be proportional to the square root of the transaction size; e.g., Gabaix et al. [2006], Bouchaud et al. [2008], Hasbrouck and Seppi [2001], and Toth et al. [2011].

[4] Examples include equities markets Hasbrouck and Saar [2009]; corporate bond markets Mahanti et al. [2008]; and interbank funding markets Gefang et al. [2011].

The literature on liquidity typically seeks to identify specific mechanisms underlying the nonlinear response. For example, Kyle and Obizhaeva [2016] derive a cube-root rule theoretically from a core assumption of Poisson arrivals of speculative order flow. A recent paper by Bookstaber et al. [2015] points to asymmetries in decision-response times between buyers and sellers as a possible source of nonlinearities in price impact.[5] Duffie [2010] suggests three explanations for the various delays across markets in the arrival of a countermanding response to an initial order-flow impulse: search, dealer capital constraints, and investor inattention. In extreme cases, a large initial price move may repel, rather than attract, price-stabilizing speculative order flow. DeLong et al. [1990] present a key early model of such positive-feedback trading. Models in this tradition are similar in spirit to the "momentum" trading explanation of Jegadeesh and Titman [1993, 2001], in that the driving force for current trading behavior is the recent history of prices alone.

Nonlinearity in liquidity is an even greater worry at the systemic level, where the stakes are correspondingly higher. Here, interactions among participants can conspire to produce self-amplifying feedback loops. Tirole [2011] provides a tour of systemic pathologies related to illiquidity, including contagion, fire sales, and market freezes. Brunnermeier [2009] describes how these forces played out during the most severe phases of the recent crisis. He highlights four channels, all of which involve liquidity: (a) deleveraging spirals driven by erosion in capital levels and increases in lending standards and margins; (b) a credit crunch motivated by funding concerns; (c) outright runs, exemplified by Bear Stearns and Lehman Brothers; and (d) netting failures due to real or perceived counterparty credit risks. We address the challenge of nonlinearity by agnostically allowing the data determine the correct number of liquidity states for each time series. Notably, for all 28 of our univariate

[5] As a possible example of this mechanism, they point to the 1987 market crash, in which relatively speedy portfolio insurance traders in the index futures markets overwhelmed the order-flow capacity of traditional equities dealers as program traders laid off the inventory in the spot market. The empirical fact that the bulk of price discovery for traded equity indexes occurs in the futures, not the spot markets, has long been recognized. See, for example, Hasbrouck [2003] and Kawaller et al. [1987].

series, three liquidity states are adequate to explain the observed variation in the price-impact statistics. Because the expected price impact is allowed to vary idiosyncratically for each of the three regimes (high, medium, and low price-impact), this model naturally captures nonlinearities in price impact.

Endogeneity means that liquidity is partly a network externality in the sense of Pagano [1989] and Economides [1996]: investors are naturally more willing to enter markets that are already very active, because of the implicit assurance that counterparties will be easy to find when needed. A familiar example is the contrast between trading for on-the-run and off-the-run Treasuries; see Barclay et al. [2006]. Similarly, Bessembinder et al. [2006] find that liquidity externalities are consistent with the significantly reduced trade execution costs that followed the introduction of the TRACE feed, which increased transparency in the corporate bond market. Liquidity externalities operate at the level of the system as well, as discussed by Morris and Shin [2004], Dang et al. [2010], and Adrian and Shin [2010], among others. We address endogeneity by estimating a hierarchical model to search for common liquidity structure throughout the cross-section of observed price-impact series. We are able to identify significant patterns and attribute them statistically to particular systemwide market summary time series, providing some economic interpretation for the estimation.

2 Model Description

This section describes our choice of a market liquidity metric, the MCMC analysis to estimate latent liquidity states, and the aggregation across markets to detect systemwide patterns.

2.1 Market Liquidity Measures

A central goal of this research is to identify broad patterns or commonalities in market liquidity that might support a formal program for monitoring systemwide liquidity conditions. This implies a difference in scope from earlier studies of liquidity commonality. By casting a wide net across diverse instrument types, we hope to have a better chance of detecting emerging liquidity anomalies and identifying key liquidity indicators and important patterns among the markets being monitored. It is impossible to know with certainty ex ante which market(s) might participate in a salient way in a systemic illiquidity event. Therefore, the cross-section of asset markets should be both broad and extensible. To be responsive to evolving liquidity conditions as a systemic surveillance tool, the liquidity measure should be available at (minimally) a daily frequency.[6] These considerations translate into four minimal requirements for an acceptable market liquidity metric:

- Feasibility – The data inputs needed to calculate the metric should be available for a broad range of markets.
- Timeliness – It should be practical to update the metric with at least a daily frequency for all markets in the sample.
- Comparability – The metric should have the same general statistical characteristics (e.g., scale and dimension) for all markets to which it is applied, to support comparisons and aggregation across a broad range of markets.
- Granularity – Measurements should be resolvable to the level of individual markets, to support attribution of systemic liquidity events to specific sectors or markets.

[6] Andersen et al. [2016] successfully extend the invariant price-impact approach to intradaily frequencies.

These criteria narrow the field of candidate measures from the research literature considerably.[7] As a practical matter, the feasibility requirement restricts attention to metrics that depend only on prices and volumes (or derived values, such as returns and volatility), because most markets have post-trade transparency of this information. In contrast, metrics requiring pre-trade transparency (e.g., quoted bid-ask spreads), customer order flow, or dealer inventories are not feasible by this definition.[8] Conditional on satisfying the feasibility requirement, the timeliness requirement usually does not bind. Even metrics requiring a multi-day estimation interval, such as the regression model for Kyle's lambda, can employ a rolling window to produce daily liquidity observations.

The comparability requirement is crucial for a systemwide analysis. Most metrics treat liquidity in one market at a time (thus providing granularity). In contrast, systemic monitoring requires the ability to understand not only whether illiquidity in a given market is unusual relative to its *own* history, but also relative to conditions in *other* markets. For many metrics, comparability might be achieved by an appropriate market-specific normalization, but other metrics are more problematic in this regard. For example, turnover (trading volume divided by total outstanding) is a commonplace heuristic for liquidity in many markets, but its interpretation differs across markets. For bonds and equities, the denominator is simply the total amount issued; for exchange-traded futures, which lack a fixed issue amount, one might substitute open interest; but for over-the-counter markets, such as interest-rate swaps or foreign exchange, the choices of a plausible denominator and interpretation of the resulting measure are ambiguous.

[7] Foucault et al. [2013], Holden et al. [2014], and Gabrielsen et al. [2011] provide a universe of market liquidity metrics to choose from.

[8] Although the magnitude of customer order flow is seldom directly observable, a number of important models, including the price-impact measure of Kyle [1985] and the $VNET$ model of Engle and Lange [2001], require only a direction-of-trade indicator, which can be inferred (with error) from the sign of sequential price changes.

2.1.1 Market-invariant price impact

Among the candidate metrics, we select the invariant price-impact measure of Kyle and Obizhaeva [2016], which captures the change in market prices caused by a one-directional order flow (buy or sell) of a given size. This metric is both feasible and timely, and is also designed to support comparability and granularity. Equation 1 shows the reduced-form (and empirically calibrated) price-impact relationship for market i.[9]

$$C_i(X_i) = \frac{\sigma_i}{0.02}\left[\frac{8.21}{10^4}\left(\frac{W_i}{W^*}\right)^{-1/3} + \frac{2.50}{10^4}\left(\frac{W_i}{W^*}\right)^{1/3}\frac{X_i}{(0.01)V_i}\right] \tag{1}$$

Equation (1) measures price impact as the market-specific (i.e., granular) expected daily volatility of returns, σ_i, normalized by a complicated scale factor (in square brackets) that adjusts for local price-level and expected volume conditions. The normalization yields a measure that should be invariant (i.e., comparable) across markets and over time. Here, $C_i(X_i)$ is the trading cost as a response to a trading impulse of size X_i, where V_i is expected daily trading volume (in shares or analogous units), X_i is a typical order size for market i, and W_i as the level of speculative activity (measured as price times expected volatility times expected volume). The first term inside square brackets is the portion of trading cost attributable to the bid-ask spread, and the second term (involving X_i) is the price-impact

[9] Kyle and Obizhaeva [2016, equations (37) and (38)] present two alternative versions of the price-impact measure, which differ in the functional form of the response of transaction costs to speculative order flow, which is allowed to be either linear, as in equation (1), or obey a square-root rule:

$$C_i(X_i) = \frac{\sigma_i}{0.02}\left[\frac{2.08}{10^4}\left(\frac{W_i}{W^*}\right)^{-1/3} + \frac{12.08}{10^4}\left(\frac{X_i}{(0.01)V_i}\right)^{1/2}\right]$$

In our sample, the two versions produce qualitatively similar results, and we focus on the linear specification.

An alternative to full cost function in equation 1 would be to use the asset-specific illiquidity measure that underlies the cost function, given as Kyle and Obizhaeva [2016, equation (12)]. This would have the benefit of focusing more narrowly on the liquidity characteristics of the asset, but would omit other influences, such as the bid-ask spread factor in the cost function, that should be relevant to systemwide liquidity. Ultimately, this is an empirical question, and a comparison of the two approaches remains as a topic for future research.

component. To facilitate an intuitive interpretation of the final result, W_i is scaled by a factor $W^* = (.02)(40)(10^6)$, which is simply a benchmark value for a hypothetical security. Although equation (1) is stated as a function of X_i, we apply a fixed rule to choose a specific value for X_i — one percent of average daily trading volume over the preceding month (20 trading days) — so that, in our implementation, equation (1) resolves to a measure based only on price, expected volume, and expected volatility.

Significantly, the normalization factor in (1) embeds significant structure, derived from theoretical first principles asserted to describe trading in speculative markets. The basic intuition of the invariance measure is that illiquidity reveals a market's resilience, or lack thereof, to net speculative order flow. Speculative bets represent individual decisions to take on (or unload) risk; they tend to arrive at different rates in different markets, creating a phenomenon of market-specific "business time" defined by the pace of speculative trading.[10] Such bets reflect the market's net risk-bearing capacity — long and short — which is the ultimate source of liquidity. Kyle and Obizhaeva [2016] argue that the bet arrivals can be approximated by a Poisson process with arrival rate γ_i, so that *expected* speculative order flow is proportional to calendar time (one unit of business time equals $1/\gamma_i$). Similarly, the observed returns variance, $\bar{\sigma}_i{}^2$, can differ across markets for many reasons, but is assumed to be a constant multiple of an underlying, market-specific betting variance, σ_i^2, for example, that caused by speculative order flow rather than news-induced volatility.

The invariance hypothesis is that, after normalizing by the local speculative capacity of the market — the amount of risk transferred per unit of business time — the arrival of a bet of dollar size PQ will generate a dollar price-impact distribution whose variance depends only on the market-specific volatility conditions: $PQ(\bar{\sigma}_i{}^2/\gamma_i)$, or, equivalently in terms

[10] This is a recent addition to the literature on "time deformation" approaches to improving the empirical regularity of financial time-series, which stretches back at least to the pioneering work of Müller et al. [1990] on intradaily foreign exchange data. Other examples include Drost and Nijman [1993], and the ACD-GARCH model of Engle [2000]. The theoretical model of Easley and O'Hara [1992] envisions a subdivision of the trading day into equally spaced (in calendar time) intervals, but with an interval length that can vary across markets to accommodate local conditions.

of standard deviation, $PQ(\bar{\sigma}_i \gamma_i^{-1/2})$. In other words, in equilibrium, overall speculative capacity allocates itself across the system to maintain an empirical relationship that is constant across markets (and over time).

Equation (1) is the final reduced-form: simple algebraic manipulations; the inclusion of a bid-ask spread component to expand the pure price-impact relationship to a more general transactions cost function, $C(X_i)$; and calibration of the remaining unidentified parameters against actual data. A factor of $\gamma_i^{3/2}$ emerges naturally in these transformations as a product of a linear scaling by business time from individual bets to observed daily volume and the square-root scaling of the volatility of the price impact distribution. This $\gamma_i^{3/2}$ factor is the ultimate source of the curious exponents in equation (1).

In applying equation (1), there is no unambiguously right way to set the typical order size, X_i. An important consideration is to normalize X_i by trading activity in each market, to measure price-impact responses on a comparable scale across markets. A corollary requirement is to calibrate equation (1) to be consistent with the chosen definition of X_i. The particular calibration in equation (1) assumes that order size is a constant fraction of average daily volume. This implies, for example, that the dollar size of the order should move in the same direction as dollar volume. A plausible alternate is to hold the dollar size of an order constant over time, so that the relative size of the order (as a fraction of volume) moves inversely with volume. As a robustness check, without re-estimating the parameters, we recalibrated order size as a constant dollar value. The price-impact results were similar in magnitude, but noisier than for the calibration of X_i as a constant fraction of daily volume; the results presented below use the constant-fraction specification.[11]

[11] Another possibility is to allow the size of the orders to adjust to market liquidity changes, in a manner more rigorously consistent with the equilibrium arguments in Kyle and Obizhaeva [2016]. For example, speculative order flow should be directly proportional to both overall liquidity and to the cube root of expected dollar volume, per Kyle and Obizhaeva [2016, equation (8)], so that order size should increase with liquidity, while decreasing as a fraction of daily volume. We are grateful to Pete Kyle for a helpful clarifying conversation around this issue. Because we are interested here in applying rather than testing or extending the Kyle and Obizhaeva [2016] model, we restrict attention to the calibrations in equations

We follow Kyle and Obizhaeva [2016] in using the average trading volume in market i over the preceding month (20 trading days) as a proxy for expected volume, $V_i = \sum_{t=-20}^{-1} v_{i,t}$, in equation (1). We similarly proxy expected volatility, $\sigma_i = \sum_{t=-20}^{-1} \sigma_{i,t}$ as the average realized volatility of daily returns over the preceding month.[12] For futures contracts, we treat each monthly maturity as a distinct market. For bonds and equities, each "market" is a value-weighted portfolio by one-digit Standard Industrial Classification (SIC) code (for equities) or ratings grade (for bonds: prime, other investment grade, high-yield, and unrated). For bonds in particular, aggregation into portfolios helps by sharply reducing the problem of missing observations due to infrequent trading for individual bonds. For each market, we calculate a daily time series of price impacts according to equation (1), using a trading impulse, X_i, set as 1 percent of the moving-average trading volume for the market. The resulting daily panel of price impact time-series across markets for futures, bonds, and equities forms the basis for our subsequent analysis.

2.1.2 Comparison to Other Approaches

The invariance measure of Kyle and Obizhaeva [2016] is only one of several established market liquidity metrics that satisfies the four requirements set out above. To further justify our choice, we compare the invariance measure (labled INVL below) to a selection of other metrics that are acceptable under our four criteria. In addition to satisfying the requirements, these alternatives were chosen to represent a diverse range of approaches:

- AMIH – This measure, defined by Amihud [2002] is based on the notion, originally advanced by [Amihud and Mendelson, 1986], that illiquidity should be priced and therefore should appear in returns. The basic equation is the daily absolute return,

(1), together with a fixed dollar size for the order impulse, X_i.

[12] There are more sophisticated ways to estimate conditional expected volatility, including the many GARCH specifications; see Bollerslev [2010]. However, Kyle and Obizhaeva [2016, p. 31] note that using a more exacting ARIMA model in log volatilities produces quantitatively similar results. Again, because we are interested in applying rather than refining their model, we apply the simpler specification.

$|R_{i,t}|$, for security i, divided by daily volume, $v_{i,t}$:

$$\text{AMIH} = \frac{|R_{i,t}|}{v_{i,t}}$$

The NYU Volatility Institute (V-Lab) [2014] at New York University provides calculated time series for three variants of the Amihud [2002] measure.

- LVOL – This is simply the logarithm of expected volatility, $\hat{\sigma}_{i,t}$:

$$\text{LVOL} = \ln(\hat{\sigma}_{i,t}),$$

  where $\hat{\sigma}_{i,t}$ is estimated as the standard deviation of daily returns over a rolling window ending at day t. We include this measure, because $\hat{\sigma}_{i,t}$ plays such a prominent role as the leading term of equation 1.

- ROLL – Roll [1984] proposes to infer (approximately) the quoted spread from the assumption that the time-series of price changes is dominated by bid-ask bounce:

$$\text{ROLL} = 2\sqrt{-\text{cov}(\Delta p_{i,t-1}, \Delta p_{i,t})}$$

  This is a workaround for the lack of pre-trade transparency in many markets.

- KLAM – Kyle's lambda, originally defined by Kyle [1985], is a commonly used price-impact measure. We calculate it as a cross-sectional average (across N firms) of estimated price-impact coefficients, $\hat{\lambda}_i$:

$$\text{KLAM} = \frac{\sum_{i=1}^{N} \hat{\lambda}_i}{N}$$

where the $\hat{\lambda}_i$ values are calculated by regressing daily returns on signed dollar volume:

$$R_{i,t} = \hat{c}_i + \hat{\lambda}_i \cdot \mathrm{Sgn}(t)\log(v_{i,t}p_{i,t}) + \epsilon_{i,t}$$

where $\mathrm{Sgn}(t) \in \{+1, -1\}$ is the direction of trade (customer buy or sell, respectively), and $R_{i,t}$, $v_{i,t}$, and $p_{i,t}$ are the return, volume and price of security i on day t.

We compare the four measures with the market invariant measure in equation 1, by estimating all five on those members of the full universe of CRSP equities with SIC code 6 (financial firms), between January 1986 and March 2014. We work with equities alone for this exercise, to reveal the behavior of the metrics over a longer time span. Large-scale systemic liquidity events like the crisis episode of 2008 are rare, and one of the purported advantages of the invariant approach is its comparability over time.[13] Understanding how the liquidity measures perform across diverse historical episodes is therefore an important exercise.

Table 1 presents simple linear correlations among the five series. Figure 1 plots the five time series over nearly 30 years, 1986–2014. Basic sample statistics appear in Table 2. All five series respond strongly to the major liquidity events of 2008, and all are positively correlated with each other, suggesting that they all indeed track some facet of market liquidity. The invariance metric (INVL) is strongly correlated with each of the four other measures, but in no case is the correlation perfect. Interestingly, the strongest correlation is between Kyle's lambda (KLAM) and log volatility (LVOL).

[13] Among the series in our sample, oil futures also have a long history, and the analysis of those series (as well as the other equities series) yields qualitatively similar results. The invariance measure registers a spike in illiquidity around the start of the first Gulf War in 1991, which does not show up significantly in equities markets. Kyle and Obizhaeva [2013] apply the invariant approach to five historical large liquidity events: the October 1929 stock market crash; the "Black Monday" crash of October 1987; a subsequent event in the futures markets, three days after the 1987 crash; the Société Générale rogue trader event in January, 2008; and the May 2010 "flash crash" in the futures markets.

Table 1: Correlations Among Liquidity Measures for Financial Equities

January 1986 – March 2014

	AMIH	LVOL	ROLL	KLAM	INVL
AMIH	1.00	0.39	0.16	0.39	0.42
LVOL		1.00	0.23	0.92	0.75
ROLL			1.00	0.11	0.44
KLAM				1.00	0.66
INVL					1.00

Sources: Center for Research in Securities Prices, Wharton Research Data Services, authors' analysis

To facilitate visual comparison, Figure 1 normalizes each series to range between zero and one. The comparison supports the Kyle and Obizhaeva [2016] metric (INVL) as an approach to systemic liquidity monitoring. The most obvious pattern in Figure 1 is the pronounced secular trend in Kyle's lambda (KLAM) and volatility (LVOL). These measures are highly correlated, due to the regression equation that creates the $\hat{\lambda}$ series, and so both appear to be picking up the "Great Moderation." From a monitoring perspective, this instability in the measures is troubling, because it complicates comparisons across liquidity events that are separated in time, and makes interpretation of the signals context-dependent. The relative lack of skewness and kurtosis for these measures (see Table 2) is also an artifact of this trend. Due to the nonlinear nature of illiquidity, one expects the metrics to be skewed. On the other hand, a visual inspection of the series reveals that the AMIH and ROLL measures are relatively noisy, with transient spikes that can be very large in scale. The other three metrics (KLAM, LVOL, and INVL) all incorporate some form of moving-window estimation that helps smooth the series. The amplified noise-to-signal ratio reduces the usefulness of the AMIH and ROLL measures as monitoring tools.

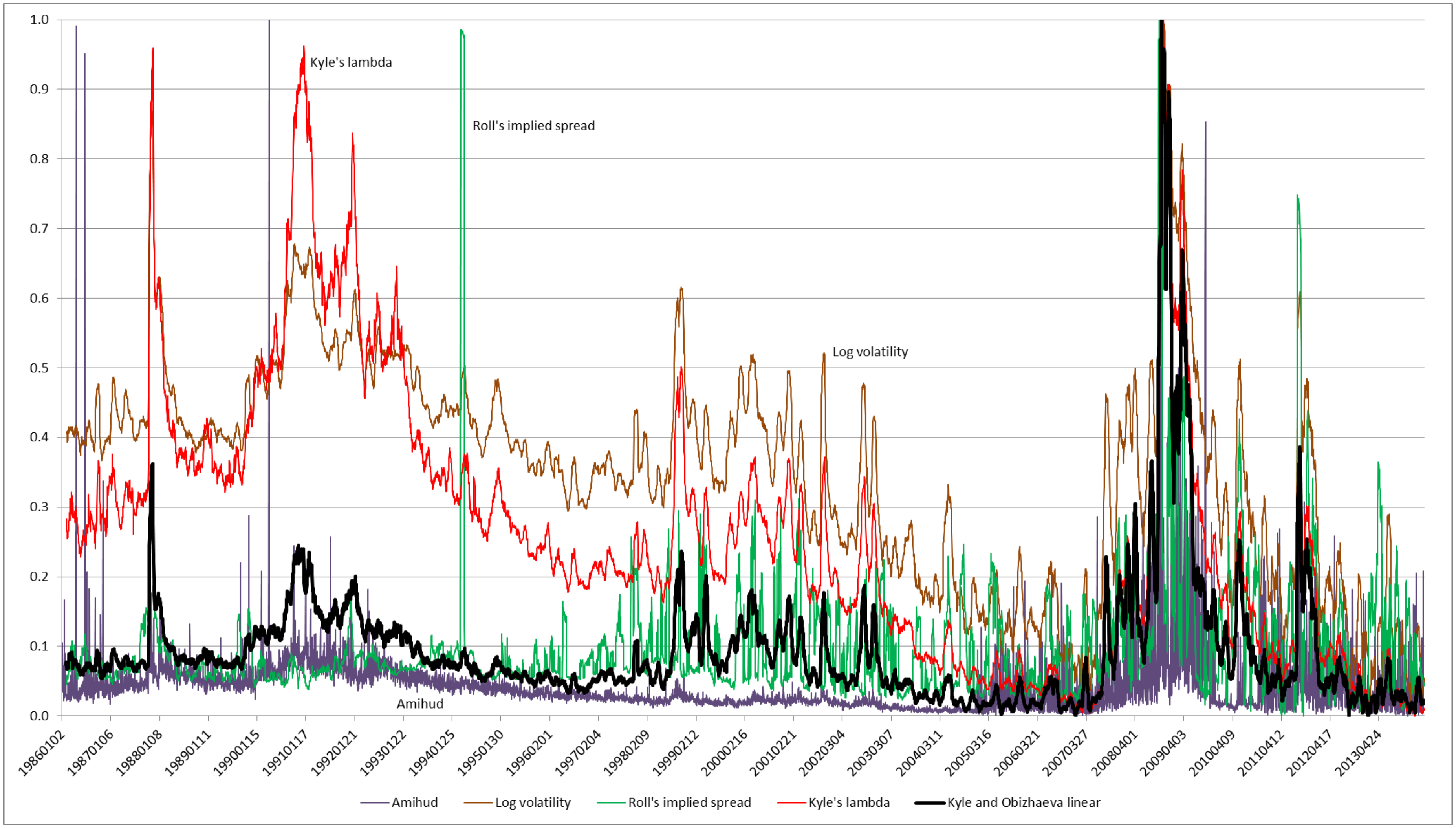


Figure 1: Five Market Liquidity Measures for SIC 6 Financial Industry Equities, January 1986 – March 2014,
Sources: Center for Research in Securities Prices, Bloomberg L.P., Mergent Inc., Wharton Research Data Services, St. Louis Federal Reserve Economic Data, authors' analysis

Table 2: Five Market Liquidity Measures for Financial Equities

January 1986 – March 2014					
	AMIH	LVOL	ROLL	KLAM	INVL
Mean	0.0433	0.3572	0.1024	0.2681	0.0971
Std. deviation	0.0480	0.1612	0.0975	0.1936	0.1005
Skewness	7.3703	0.2686	4.9402	1.0985	4.3027
Kurtosis	101.5915	0.5276	35.0150	1.2380	25.5954
Num. observations	7121	7121	7121	7121	7121

Sources: Center for Research in Securities Prices, Wharton Research Data Services, authors' analysis

Overall, the comparison supports the selection of the invariant price-impact measure of Kyle and Obizhaeva [2016] for systemwide monitoring of market liquidity conditions. Most importantly, it satisfies the requirements of feasibility, timeliness, comparability, and granularity set out at the beginning of this section. In addition, the invariance measure exhibits more consistency over time and is less subject to general noise and transient spikes, compared to several other commonly used market liquidity metrics.

2.2 Univariate Models of Latent Structure

This section describes the hidden Markov chain (HMC) estimation of latent structure in the price-impact series. The primary assumption underlying the HMC approach is that the liquidity in a given market jumps between distinct states, such as from low to high liquidity, and then stays in that new state for a random period of time. The observed level of liquidity is then a random deviation from the average liquidity level associated with this underlying state. Although observed liquidity can occasionally fall between the average liquidity levels for two neighboring states, the persistence of the observed liquidity allows the HMC model to resolve ambiguity regarding identification of the liquidity state at any particular point in time.

When considering liquidity across multiple financial markets (each possibly represented by a portfolio of securities), we augment the univariate liquidity models for the individual markets with an add-on hierarchical model. This approach helps us explain, in part, periods of coordination where the broader financial system exhibits similar liquidity patterns across the individual markets (see subsection 2.3 below). The hierarchical model is an "add on" in the sense that the aggregate patterns it summarizes do not feed back into the underlying univariate HMC models. Instead, the add-on model helps reveal the extent to which a collection of systemwide economic summaries (*e.g.*, TED spread, Dow Jones U.S. Real Estate Index, etc.) explain the relationship between general economic conditions and liquidity states across the entire financial system. Uncovering these relationships offers a framework for understanding — and potentially predicting — systemwide liquidity stress, either by lagging the summary series and/or predicting the underlying dynamics of these systemwide economic summaries.

We focus on regime-shifting models, because they naturally account for sudden jumps in liquidity, which are symptomatic of stress in the broader financial markets. A regime-shifting model also allows us to remove slow-moving diffusion dynamics from our hierarchical model, which, for the data that we consider, improves our predictive ability. As a robustness check on our add-on hierarchical model approach, we explored a traditional vector autoregressive (VAR) model, but found that VAR models tend to reduce to individual autoregressive models, where each price-impact series is driven primarily by its own lagged price impact.[14] Even when we include the systemwide economic summaries, the VAR models offer limited insights and almost no predictive ability.

In an acknowledgment that the liquidity series sometimes exhibit slow-moving diffusion dynamics alongside jumps, we introduce a version of the univariate model that allows for autoregression around the average level of attraction for each liquidity state.[15]

[14]See Koop and Korobilis [2010] for a discussion of Bayesian approach to VAR models. We discuss our VAR results briefly in Section 5.

[15]An alternative, ad hoc approach would be to define regimes by breaking the liquidity series into

For each market, we propose a univariate HMC model where measured price impact is a random deviation from some latent level of price impact associated with a state of the underlying HMC. We consider two types of random deviations: i) independent deviations around an average price impact, and ii) independent deviations around a value that mean-reverts around an average price impact level. Initially, we assume the dynamics of the individual univariate HMC models are unrelated. Section 2.3 introduces the hierarchical, add-on model that uses systemwide economic summaries to explain how the hidden states, identified by the underlying collection of univariate hidden Markov models, tend to evolve.

Both of the univariate HMC models assume that the liquidity measurements over T periods for market i, $y_i = (y_{i1}, ..., y_{iT})^\intercal$, are independent, normally distributed deviations around a dynamic, latent level of price impact, $\theta_i = (\theta_{i1}, ..., \theta_{iT})$:

$$y_i = \theta_i + \epsilon_i,$$

where $\epsilon_i \sim N(0, \sigma_i^2 I_T)$, and I_T is a T-dimensional identity matrix. For the first version of the model (HMC-only), each of the K elements of θ_i represents a different liquidity state for each market-specific, discrete-time HMC, D_i. We map the latent level of attraction for each state, k, for market i as a function of the HMC, D_i, as follows:

$$\theta_i = F_i(t, k)\bar{\theta}_i = F_i\bar{\theta}_i,$$

where $\bar{\theta}_i$ is a $K \times 1$ vector, with each element of $\bar{\theta}_i$ representing the average level of price

percentile regions, such as lower quartile, interior quartiles, and top quartile. This approach, however, would not take advantage of the persistence built into a HMC model. As a result, it would tend to produce an estimate of states with frequent jumps, even if these changes in price impact are spurious. More importantly, this data summary approach would not find the long-term levels of attraction that exist naturally in the data.

impact associated with the k^{th} state of D_i, and

$$F_i(t,k) = I\{D_{it} = k\},$$

where $I\{\}$ is an indicator function is 1 if the argument is true and 0 otherwise.

The HMC version of the model is typically sufficient to identify structural shifts in liquidity patterns. However, there are some markets for which the local variation in the level of liquidity supports an excessive number of hidden states. In these cases, we use a mean-reverting version of the hidden Markov chain model. For this mean-reverting hidden Markov chain (MRHMC), the latent level θ_i mean-reverts around an average level associated with each of the state of D_i, or for $t = 2, ..., T$,

$$\Delta\theta_i = \gamma_i((\theta_i)_{\setminus T} - (F_i\bar{\theta}_i)_{\setminus 1}) + (\xi_i)_{\setminus 1}, \tag{2}$$

where the subscript notation $(x)_{\setminus 1}$ indicates that the first element has been removed and $(x)_{\setminus T}$ indicates that the last element of the vector (x) has been removed, and $\Delta\theta_{it} = \theta_{it} - \theta_{it-1}$. For $t = 1$ let

$$\gamma_i\theta_{i1} = \gamma_i\bar{\theta}_{i1} + \xi_{i1},$$

where $\xi_i \sim N(0, w_i I_T)$. We require $0 < \gamma_i \leq 1$, which ensures that θ_i is stationary and increases the variance of θ_{i1}, allowing the starting value of θ_i to be relatively vague. Alternatively, we can rewrite (2) as:

$$L_i\theta_i = \gamma_i F\bar{\theta}_i + \xi_i, \tag{3}$$

where L_i is a sparse $T \times T$ matrix with zeros except for the following elements, $L_i(j,j) = 1$

and $L_i(j, j-1) = \gamma_i - 1$ for $j > 1$ and $L_i(1,1) = \gamma_i$.

For both versions of the univariate HMC model, the dynamics of D_i are given by an initial probability density, ν_i (which is a $K \times 1$ vector), and the corresponding $K \times K$ transition probability density, P_i. Given a realization of D_i, the density of the HMC is given by

$$f(D_i) = \nu(D_{i0}) \prod_{t=1}^{T} P_i(D_{it-1}, D_{it}).$$

We assume conjugate priors for σ_i^2 and w_i (inverted Gamma), ν_i and each row of P_i (Dirichlet) and $\bar{\theta}_i$ and γ_i (truncated normal). In addition, for the MRHMC version of the model we use subjective priors based on initial, conditional maximum likelihood estimates of parameters based on various summaries of the data, to ensure that the filtered HMC model can clearly distinguish between the dynamics of the hidden Markov chain and the dynamics of the latent value θ_i.

2.2.1 Full Conditional Distributions HMC Model

We use Markov chain Monte Carlo (MCMC) analysis to infer parameter values for both the univariate models as well as the hierarchical model we build on top of them.[16] The full conditional densities used in the MCMC analysis for the HMC model are:

$$\bar{\theta}_i|- \sim N\left(\left(\frac{1}{\sigma_i^2}F_i^T F_i + \frac{1}{\tau_{\bar{\theta}_i}^2} I_K\right)^{-1}\left(\frac{1}{\sigma_i^2}F_i^T y_i + \frac{\mu_{\bar{\theta}_i}}{\tau_{\bar{\theta}_i}^2}\right), \left(\frac{1}{\sigma_i^2}F_i^T F_i + \frac{1}{\tau_{\bar{\theta}_i}} I_K\right)^{-1}\right) I\{\bar{\theta}_{i1} < ... < \bar{\theta}_{iK}\};$$

where $-$ represents all other conditioning information remaining in the model, and

$$\frac{1}{\sigma_i^2}|- \sim Gamma\left(shape_{\sigma_i^2} + \frac{T}{2}, scale_{\sigma_i^2} + \frac{1}{2}(y_i - F_i\bar{\theta}_i)^T(y_i - F_i\bar{\theta}_i)\right).$$

[16]For a description of MCMC methods, see Brooks et al. [2011] and Gelman et al. [2013].

Realizations of the hidden Markov chain D_i, conditional on the remaining parameters and data, are generated following the filter-forward, sample-backward approach commonly used with discrete-time Hidden Markov chains.[17] For completeness, the filter-forward equations for the HMC model are given by:

$$f\left(y_{it}|-,\mathcal{F}_{it-1}\right)=\sum_{k=1}^{K}f\left(y_{it}|D_{it}=k,-,\mathcal{F}_{it-1}\right)f\left(D_{it}=k|-,\mathcal{F}_{it-1}\right), \tag{4}$$

where $\mathcal{F}_{it}=\{Y_{i1},...,Y_{it}\}$; and by

$$f\left(D_{it}=k|-,\mathcal{F}_{it}\right)=\frac{f\left(y_{it}|D_{it}=k,-,\mathcal{F}_{it-1}\right)f\left(D_{it}=k|-,\mathcal{F}_{it-1}\right)}{f\left(y_{it}|-,\mathcal{F}_{it-1}\right)}. \tag{5}$$

Specifying a vague initial state probability, for example,

$$f(D_{i0}=k|-,\mathcal{F}_{i0})=\frac{1}{K},$$

completes the forward recursion. The key equation for the backward sampling is the density of the HMC conditional on all of the data, or

$$\begin{aligned} &f\left(D_{iT-t}=k|-,\mathcal{F}_{iT}\right)= \\ &\textstyle\sum_{j=1}^{K}\frac{f(D_{iT-t+1}=j|D_{iT-t}=k,-,\mathcal{F}_{iT-t})f(D_{iT-t}=k|-,\mathcal{F}_{iT-t})}{f(D_{iT-t+1}=j|-,\mathcal{F}_{iT-t})}f\left(D_{iT-t+1}=j|-,\mathcal{F}_{iT-t+1}\right). \end{aligned} \tag{6}$$

Given these formulas, it is straightforward to generate a realization of D_i: i) calculate the forward filter; ii) generate a sample for D_{iT} from (5), with $t=T$; and iii) recursively calculate $f\left(D_{iT-t}=k|-,\mathcal{F}_{iT}\right)$, conditional on all of the draws $(D_{iT},...,D_{iT-t+1})$ using (6), and then use this density to generate a sample for D_{iT-t}. Given a realization of D_i, the

[17] See Baum et al. [1970]. Cappé et al. [2005] describe these in a more general continuous-time framework.

full conditional distribution for each row of the transition probability is given by:

$$P_i(j,:)|- \sim Dirichlet\left(\alpha_{ij1} + n_{ij1}, ..., \alpha_{ijK} + n_{ijK}\right),$$

where α_{ijk} is the prior associated with D_i jumping from state j to k and n_{ijk} is the actual number of times that the current realization of D_i jumps from state j to state k. A similar full conditional density exists for ν_i, but it is inconsequential because the backward recursion dominates the *a priori* initial state. It is important to note that for the MRHMC model we are disentangling two sets of dynamics – the dynamics of D_i and θ_i. In practice, we found that the model required a strong priors on the dynamics of D_i to obtain meaningful distinctions between these two dynamics. Setting α_{ikk} to a sufficiently large value for each state k, suggests *a priori* that the hidden chain is persistent. This means that, *a priori*, once the HMC enters a state, it tends to remain in that state for a long time. This assumption results in a clean separation between the autoregressive dynamics, which are shorter in duration, and the HMC dynamics, which are longer.

2.2.2 Full Conditional Distributions and the MRHMC Model

There are similarities between some of the full conditional densities of the MRHMC model and the HMC model. The full conditional density of P_i for the HMC model is unchanged in the MRHMC model, while the full conditional density of $\bar{\theta}_i$ for the MRHMC model is obtained by replacing y_i with $\frac{1}{\gamma_i}L_i\theta_i$ and σ_i^2 with $\frac{w_i}{\gamma_i^2}$ in the corresponding full conditional density of the HMC model. The full conditional density of $\frac{1}{\sigma_i^2}$ for the MRHMC model is obtained by replacing $F_i\bar{\theta}_i$ with θ_i in the corresponding full conditional density of the HMC model, and the full conditional density of D_i for the MRHMC model is obtained by replacing the likelihood $f\left(y_{it}|D_{it} = k, -, \mathcal{F}_{it-1}\right)$ used in (4) and (5) for the full conditional density of the HMC model, with $f\left(\theta_{it}|D_{it} = k, -, \theta_{it-1}, ..., \theta_{i1}\right)$. The remaining full conditional densities for the MRHMC model are as follows:

$$\frac{1}{w_i}|- \sim Gamma\left(shape_{w_i} + \frac{T}{2}, scale_{w_i} + \frac{1}{2}(L_i\theta_i - \gamma_i F_i\bar{\theta}_i)^T(L_i\theta_i - \gamma_i F_i\bar{\theta}_i)\right);$$

$$\gamma_i|- \sim N\left(\Sigma_i\left(\frac{1}{w_i}A_i^T\Delta\theta_i + \frac{\mu_{\gamma_i}}{\tau_{\gamma_i}^2}\right), \Sigma_i\right) I\{0 < \gamma_i \leq 1\};$$

where

$$\Sigma_i = \left(\frac{1}{w_i}\left(A_i^T A_i + (\theta_{i1} - \bar{\theta}_{iD_{i1}})^2\right) + \frac{1}{\tau_{\gamma_i}^2}\right)^{-1} \text{ and } A_i = \left((\theta_i)_{-T} - (F_i\bar{\theta}_i)_{-1}\right),$$

and

$$\theta_i|- \sim N\left(\left(\frac{1}{w_i}B_i + \frac{1}{\sigma_i^2}I_K\right)^{-1}\left(\frac{\gamma_i}{w_i}B_i\left(L_i^{-1}F_i\bar{\theta}_i\right) + \frac{1}{\sigma_i^2}y_i\right), \left(\frac{1}{w_i}B_i + \frac{1}{\sigma_i^2}I_K\right)^{-1}\right),$$

where B_i is a $T \times T$ matrix given by

$$B_i = \left(L_i^{-1}(L_i^{-1})^T\right)^{-1}.$$

An alternate approach for sampling θ_i and γ_i, conditional on D_i, is to treat them as a discrete-time, dynamic linear model and use a filter-forward, sample-backward strategy like the Kalman filter.[18] Although we explored a filter-forward, sample-backward approach, we found that this was not as stable as the regression-based approach detailed above. One disadvantage of the regression approach is the need to calculate B_i, which requires the inversion of a $T \times T$ matrix. While this step may appear to be computationally prohibitive as T becomes large, the form of L_i results in a banded matrix for B_i, where all elements are

[18] See Kálmán [1960] for the original reference, and Carter and Kohn [1994] and De Jong and Shephard [1995] for MCMC-based inference approaches.

zeros except for the main diagonal and neighboring diagonals. In addition, the non-zero elements are functions of γ_i; to be explicit,

$$B_i(j,j) = \begin{cases} 1, & \text{if } j = T \\ \frac{1}{2} + 2\left(\frac{1}{2} - \gamma_i\right)^2, & \text{otherwise} \end{cases}$$

and $B_i(j, j-1) = B_i(j-1, j) = \gamma_i - 1$.

2.3 Hierarchical Model

We implement the hierarchical "add-on" model as an auxiliary analysis, performed in conjunction with the analysis of a collection of underlying HMC or MRHMC models, one for each of the N markets being considered. The add-on model treats these realizations as data and searches for patterns and relationships among these N markets. During the analysis, the N independent, underlying MCMC algorithms repeatedly generate realizations of the HMC for each market being considered, $(D_1)^n, ..., (D_N)^n$, where $(D_i)^n$ is the n^{th} realization of the HMC generated by the MCMC algorithm for the i^{th} market. In explaining systemwide liquidity regimes (see Section 4.2 below), we pool all of the HMC realizations for all of the markets to be used as data (*i.e.* the dependent variables) for a probit model. While the add-on model uses the HMC realizations as data, it is an auxiliary model in the sense that its parameters do not influence the underlying MCMC analysis of the HMC or MRHMC models generating the realizations $(D_1)^n, ..., (D_N)^n$. The goals of this auxiliary model are twofold: first, to summarize and explain patterns in the collection of hidden Markov chains; and secondly, to exploit these summaries to make predictions of future liquidity states.

The add-on model assumes that the state of each realization of the hidden Markov chain at each point in time point, D_{it}^n, is a draw from a multinomial probit distribution, which is driven by a set of systemwide summaries of the broader market, x_t.[19] These systemwide summaries are daily financial time series, listed in Table 3 and discussed in Section 3.

To be more explicit, the add-on model defines the probability that any of these N different markets will be in a different latent liquidity states as:

$$f\left(D_{it}=k\right)=f\left(\tilde{z}_{itk}>\tilde{z}_{itl},l\neq k\right), \tag{7}$$

where $\tilde{z}_{it}$ is multivariate normal:

$$\tilde{z}_{it}\sim N\left(\tilde{\beta}x_t,\tilde{\Sigma}\right).$$

We follow the approach of McCulloch and Rossi [1994], which builds on Albert and Chib [1993], to handle the identification issues that arise in a Bayesian estimation for the multinomial probit model. McCulloch et al. [2000] discuss an alternative approach. We overcome the additive identification problem by forcing the latent value for state 1 always to be zero, achieved by defining z_{it} as:

$$z_{itk}=\tilde{z}_{itk}-\tilde{z}_{it1},$$

which results in (7) becoming

$$f\left(D_{it}=k\right)=\begin{cases} f\left(0>z_{itl}l\neq k\right), & \text{if } k=1\\ f\left(z_{itk}>\max(0,z_{itl}),l\neq k,l>1\right), & \text{if } k>1\end{cases}$$

[19]Possible alternatives to the multinomial probit include an ordered probit model. The main advantage of the multinomial probit model in this setting is that it allows for increasing (or decreasing) levels of the covariates to favor a middle liquidity state, which is what we find in our data. An ordered probit model could accommodate this type of phenomena, if it included quadratic terms of the covariates. Given the behavior that we observed, we felt that the multinomial probit model offered a reasonable and parsimonious modeling approach, compared to the ordered probit model.

where $z_{it} \sim N(\beta x_t, \Sigma)$, and β is a $(K-1) \times p$ matrix, where p is the number of summary series, including an intercept. The scale identification is achieved by restricting $\Sigma_{1,1} = 1$. We assume conjugate priors for β and Σ. Following McCulloch and Rossi [1994], we sample β and Σ from the unconstrained full conditional densities using Gibb samplers and then rescale by dividing these draws by $\Sigma_{1,1}$, which enforces the above constraint.

The auxiliary model approach has the interesting property that the "data" used to calibrate the probit model — i.e., the realizations of the hidden Markov chains — can and do change values during the MCMC analysis of the underlying univariate model. Because each sweep of the N underlying MCMC analyses results in a new realization, those portions of the hidden Markov chains that are relatively stable (i.e., have a high probability of being in only one state) have the largest impact on the multinomial parameter estimates of the hierarchical model. Conversely, portions of the hidden Markov chain that tend to switch states (have a probability that is distributed between two or more states) have less impact, because the hidden Markov chains alternate between these competing states during the analysis, requiring estimates of the multinomial parameters, β, that can reasonably accommodate these oscillations. Because the hidden Markov chains tend to have a high probability of being in only one state for an overwhelming portion of the time (in our sample), this auxiliary approach gives a reasonable summary of the relationship between these systemwide summaries and the broader patterns of liquidity.

Finally, we acknowledge that a more integrated approach for the hierarchical model, going beyond the "add on" approach, would be to model the transitions of the N Markov chains from each state to all possible other states using a multinomial probit model, where these probabilities are again driven by a set of systemwide summaries. This approach would require — assuming three liquidity states — separate multinomial probit models to describe i) the probability of moving from the low liquidity state to any of the another states, ii) the probability of moving away from the medium state, and iii) the probability of moving away from the high state. It would also result in a model where the probit parameters directly affect the draws of the hidden Markov chains during the MCMC analysis. Because the "add on" approach was able both to recover the average trends of the market dynamics and had strong out-of-sample predictive performance, we have left the exploration of an integrated hierarchical model as a topic for future research.

3 Data

We measure market liquidity on a daily basis across 28 markets, covering thousands of individual securities in four different asset classes. One important goal of casting a wide net across a diverse sample is to improve the chances of identifying emerging risks in liquidity, since it is difficult to assert a priori which market sector(s) might be affected first in a systemwide episode of illiquidity. Similarly, a broad panel should help in discerning significant patterns among the markets, as we map between local markets and system-level conditions. Finally, we hope to extend our collection of distinct asset classes in future research of broader systemic liquidity conditions.

Specifically, our initial dataset includes the following markets:

- All U.S. equities, January 1986 – March 2014, from CRSP, which provides security price, return, and volume data for the NYSE, AMEX, and NASDAQ stock markets. CRSP also provides the SIC code for each security.

- All U.S. corporate bonds, July 2002 – March 2014, from TRACE, the Financial Industry Regulatory Authority's (FINRA) real-time price dissemination service for the over-the-counter bond market. It provides transaction data for all eligible corporate bonds, which include investment grade and high-yield debt; we use the public TRACE database in this analysis.[20] We map individual bonds from the TRACE

[20]We apply the heuristics of Dick-Nielsen [2009] to scrub the TRACE data. There is a separate "en-

database to their agency ratings from Moody's Investors Service, Standard & Poor's Financial Services, and Fitch Ratings, Inc., which we take from the Mergent Fixed Income Securities Datbase (FISD).

- West Texas Intermediate (WTI) light sweet crude oil futures, January 1986 – March 2014, from the New York Mercantile Exchange, the world's largest-volume futures contract traded on a physical commodity. We collected data from Bloomberg for WTI contracts with expiration maturities ranging from one to six months.

- S&P 500 market volatility index (VIX®) futures, April 2004 – March 2014, from the Chicago Board Options Exchange. This is a pure-play contract on implied volatility designed to reflect investors' view of future (30-day) expected stock market volatility. We collected data from Bloomberg for VIX® contracts with expiration maturities ranging from one to nine months.

Our primary analysis of the liquidity measures starts in 2004, when all series are available. We also provide some secondary comparisons of the longer-term performance of price-impact measures for equities and WTI futures, extending back to 1986. We grouped CRSP equities data into value-weighted portfolios based on one-digit SIC codes. The SIC portfolios cover SIC codes 0 through 8.[21] We grouped the TRACE corporate bond data into four value-weighted portfolios based on ratings grades: (1) prime, (2) non-prime investment grade, (3) speculative, and (4) unrated/unknown. Clustering into portfolios reduces the dimensionality of the analysis and presentation of results. For corporate bonds, the aggregation into portfolios is a practical necessity for the calculation of returns and volatility, because the trading of individual issues in this market is too thin.

hanced" version of the TRACE database, FINRA [2009], which does not truncate large trades, but which FINRA publishes only with a lag.

[21]The miscellaneous category (SIC 9, government establishments) is very lightly populated and did not provide sufficient observations for reliable analysis.

We track the VIX® and WTI futures at the level of their relative maturity date, starting with the front-month contract. Trading volumes by actual calendar maturities follow a sawtooth pattern, as expiry dates gradually approach and abruptly transition to the next contract as expiration occurs. For both VIX® and WTI, and for futures markets generally, near-dated contracts are usually more actively traded than the longer-maturity futures. There is no official longest maturity, but many possible long-dated contracts simply never trade. For the VIX® futures, we draw the line at nine different securities from the front month out to nine months forward. For the WTI futures, we use six different securities from the front month out to six months forward.

For the hierarchical add-on analysis discussed in Section 4.2, we use 11 daily time series, which form a panel of systemwide summaries described in Table 3. The hierarchical analysis fits multinomial probit models that use these measures of overall market conditions to explain the univariate liquidity state dynamics from the MCMC analysis.

4 Liquidity Regimes

We initially estimated each price impact series independently, using both the hidden Markov chain (HMC) and the mean reverting hidden Markov chain (MRHMC) models. Although there is no coordination between the dynamics of the latent liquidity states across markets for this initial analysis, we find surprising consistency in the dynamics of liquidity across all markets. Despite these common features, we also find interesting differences across the various markets in the lead-up to the 2008 failure of Lehman Brothers and in its aftermath. We formally explore these difference using the hierarchical model, which allows us to link the latent liquidity states from multiple markets together with a collection of systemwide summaries of the broader financial market. This provides a framework for assessing the explanatory and predictive power of these systemwide summaries.

Table 3: Summary Series for the Hierarchical (Multiple Market) Model

Variable	Description	Bloomberg Ticker
3-month Repo Rate	ICAP General Collateral Treasury 3-month repurchase agreement rate	USRGCGC ICUS
Yield Curve	Yield on the constant maturity 10-year U.S. Treasury bond minus the yield on the constant maturity 2-year U.S. Treasury note	USYC2Y10
TED Spread	3-month LIBOR rate minus the 3-month U.S. Treasury bill yield	BASPTDSP
Moody's Baa Corporate Bond Index	Yield on the Moody's investment grade long-term corporate bond index	MOODCBAA
VIX® Index	Reflects the market estimate of future (30-day) volatility of the S&P 500	VIX
Dow Jones U.S. Real Estate Index	Index of real estate investment trusts (REITs) and other companies investing directly or indirectly in real estate through development, management, or ownership	DJUSRE
S&P 500 Price-to-book Ratio	Ratio of equity market value to book value per share of the S&P 500	SPX
Three-month LIBOR-OIS Spread	Difference between the 3-month LIBOR and the 3-month U.S. dollar overnight index swap (OIS) rate	BICLOISS
5-year U.S. Breakeven Inflation Rate	Calculated by subtracting the real yield of the 5-year inflation-linked maturity curve from the yield of the closest 5-year nominal Treasury maturity. The result is the market-implied inflation expectation over the next 5 years	USGGBE05
WTI Front-month Price	Futures price for the near-dated expiry of the WTI oil contract	CL1
U.S. Dollar Index	Indicates the general international value of the U.S. dollar, by averaging exchange of the dollar against other major currencies	DXY
	Source: authors' analysis	

4.1 Individual Market Liquidity

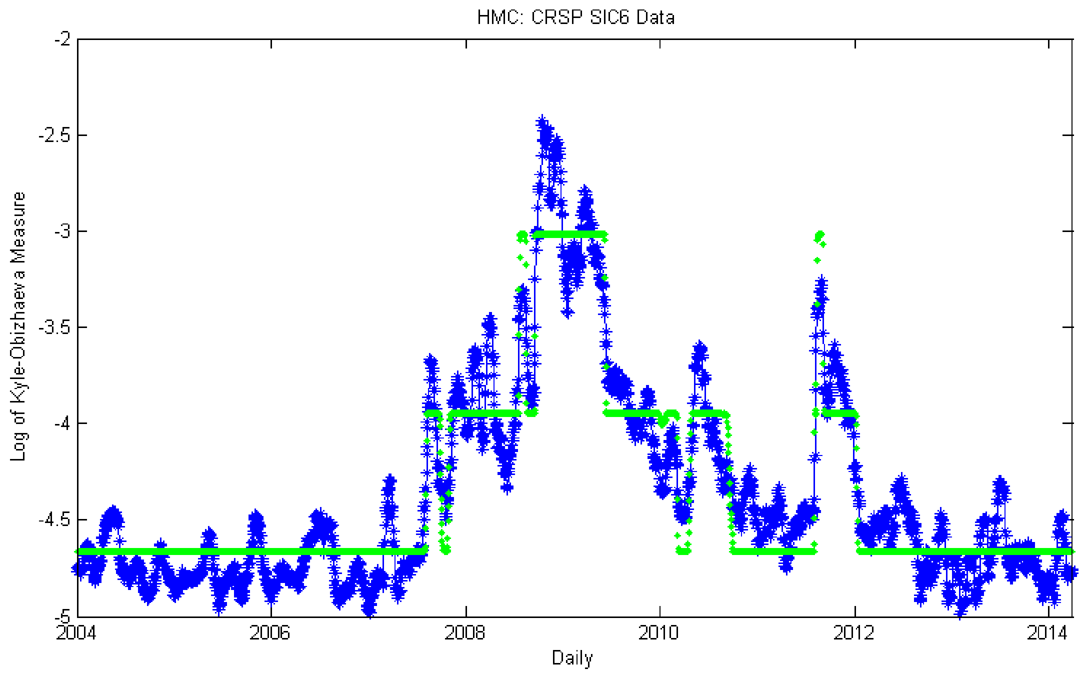
We start by considering the performance of the two competing univariate models (HMC and MRHMC) and provide evidence that there are essentially three different liquidity regimes across our sample of markets. Then we report aggregate summaries based on these models.

Performance of Models

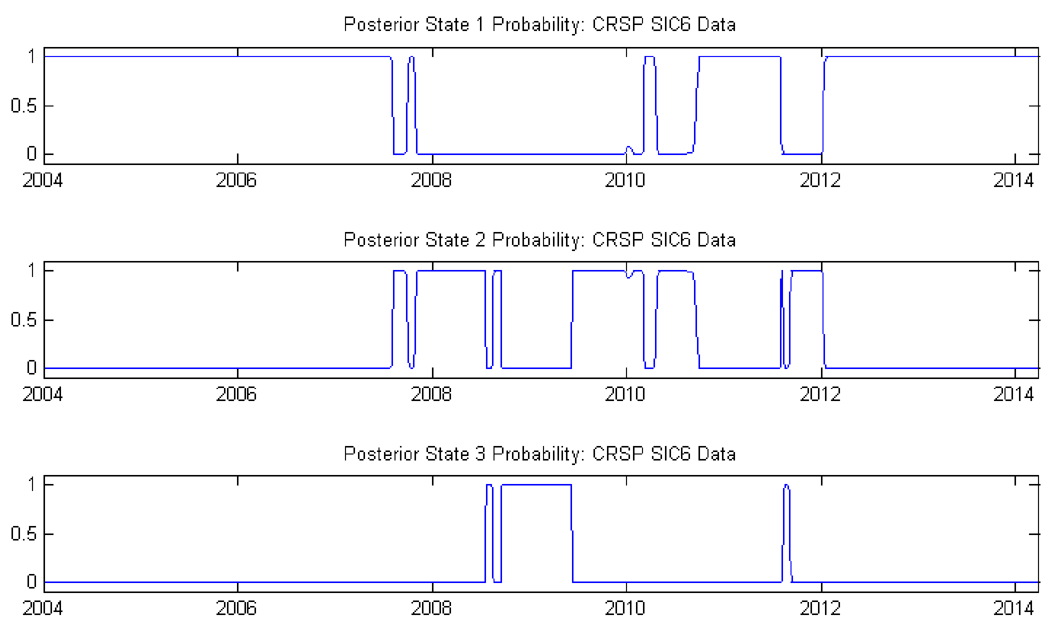
Both the HMC and MRHMC models identify interesting liquidity regimes within the Kyle-Obizhaeva price-impact data over the various markets in our sample. However, the relative performance of the two models depends on the amount of local variability of the liquidity in each state. The simpler HMC model can readily identify the three liquidity regimes for all of the equity markets; for example, Figure 2 shows results for the SIC6 equities portfolio (financial stocks). The blue time series plots the estimated daily price impact metric for the SIC6 portfolio, while the green series — nearly a step function — plots the expected daily price impact from the fitted HMC model. In some cases, such as the front month of the WTI contract (see Figure 3), the MRHMC model performed slightly better than the HMC model using standard Bayesian model-choice tools.

As a tool for policy makers, it is important to have a model that is parsimonious with respect to the number of states.[22] This implies a strong prior on the model space, giving a high penalty for increased complexity, measured as the number of hidden states. The upshot is that the fitted model prefers three states for most markets. Alternative priors, with a smaller penalty for complexity, would support a high number of latent states (as many as 10 to 15); a visual inspection indicated that this increase in the number of states tended to fracture our intermediate-liquidity state into a larger number of substates.

[22]Policy makers are the primary audience for this model, although we anticipate market participants will find it valuable as well. A framework with three states for each market, where the third state captures conditions of extreme illiquidity, has advantages for policy makers, who must make decisions in crisis conditions. Of course, it is straightforward to redo the entire analysis if policy makers or market participants feel that more or fewer states would provide a more useful insights.

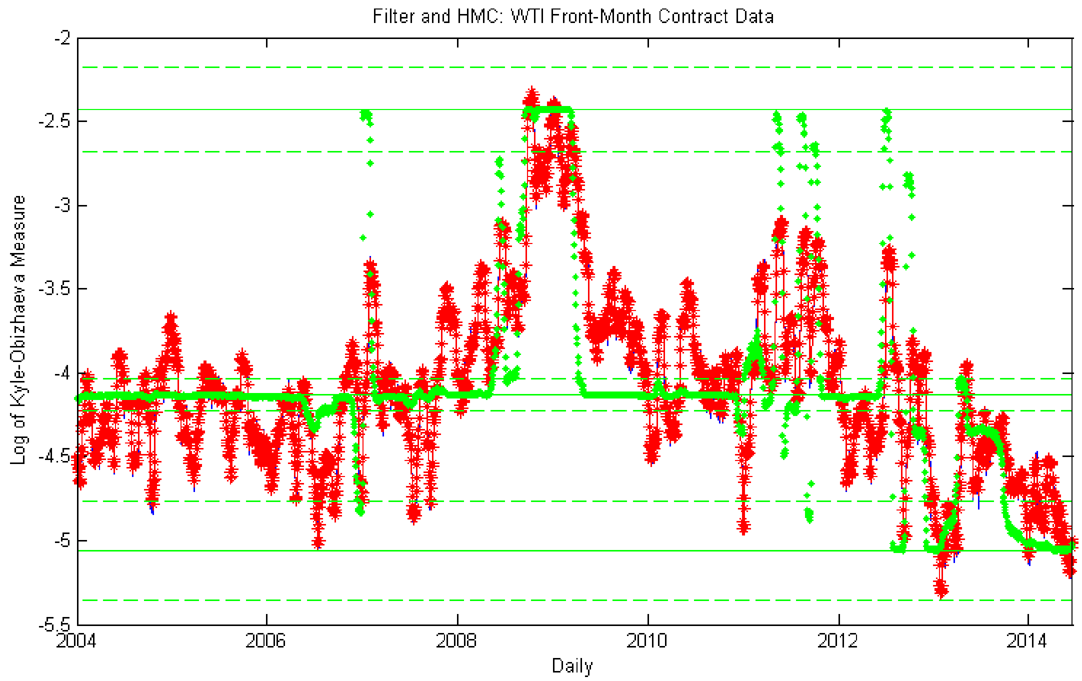


(a) Price-impact measure & level of attraction

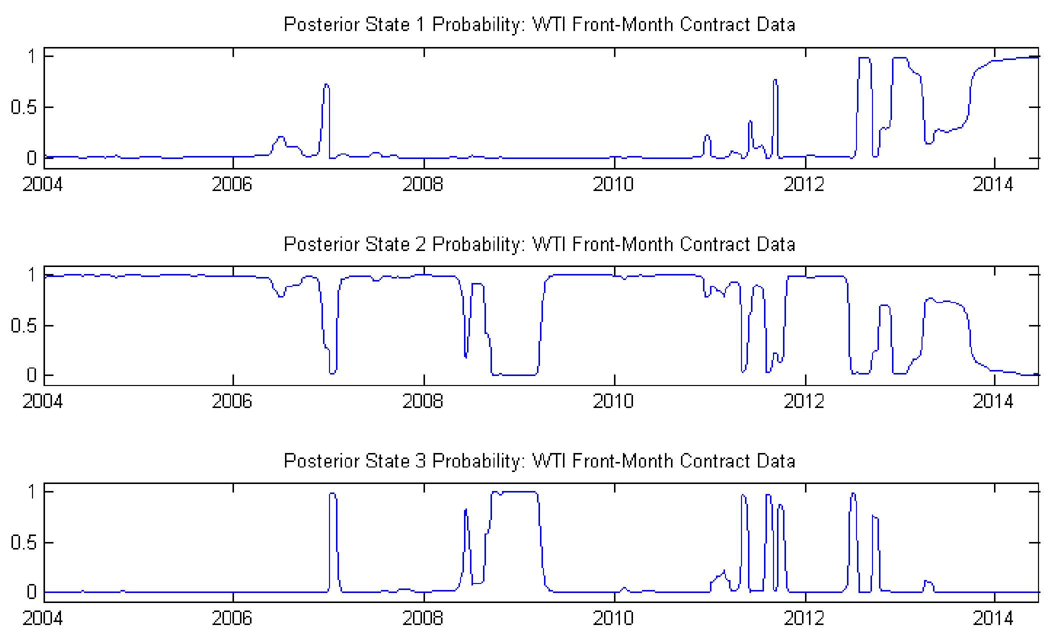


(b) HMC state probabilities

Figure 2: Equities, SIC 6, Kyle-Obizhaeva Measure and HMC Estimates
Sources: Center for Research in Securities Prices, Wharton Research Data Services, authors' analysis



(a) Price-impact measure, mean-reversion level, & level of attraction



(b) HMC state probabilities

Figure 3: WTI Futures, Front Month Contract, Kyle-Obizhaeva Measure and MRHMC Estimates
Sources: Bloomberg L.P., authors' analysis

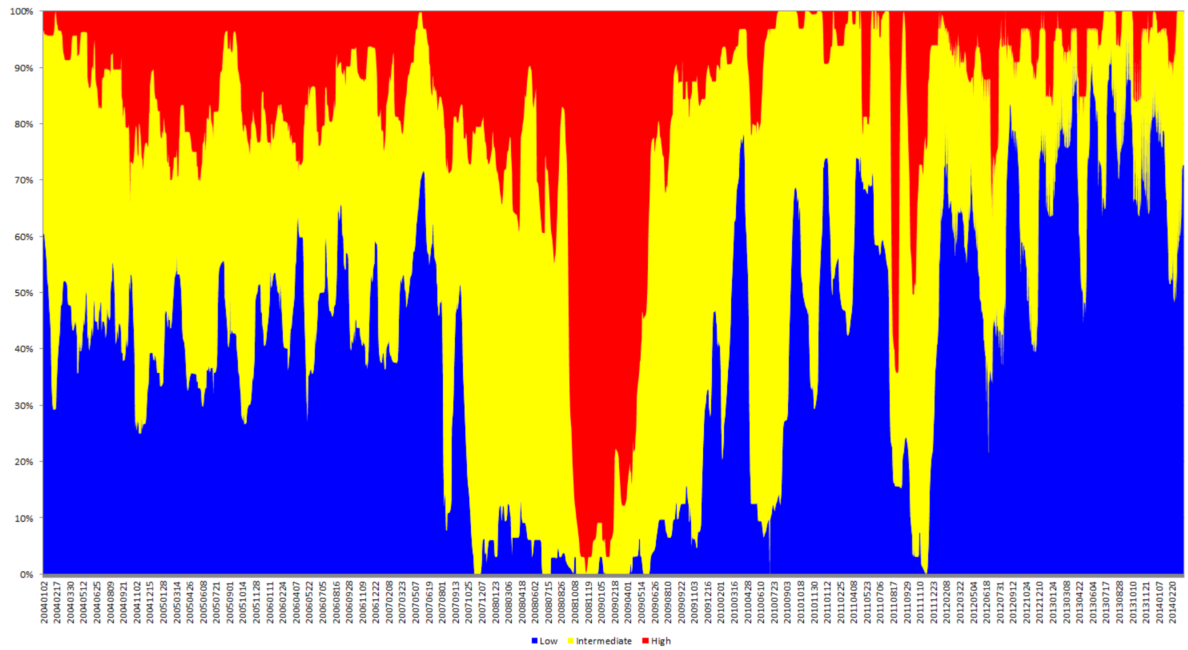


Figure 4: Probabilities of Price-Impact States, Averaged across 28 Markets, April 2004 – March 2014
Sources: Center for Research in Securities Prices, Bloomberg L.P., Mergent Inc., Wharton Research Data Services, Financial Industry Regulatory Authority, authors' analysis

For each of our 28 markets, the subsequent analysis works exclusively with a three-state HMC.[23] We label the states as the (i) low, (ii) intermediate, and (iii) high price-impact states for each series, where high price impact means low liquidity and vice versa. The analysis produces a daily estimate of the probability that each market was in each of these three unobserved states. Figure 4 presents the cross-sectional averages across the 28 series of the three probabilities of high price impact (red), low price impact (blue), and intermediate price impact (yellow). The three probabilities must sum to one each day.

While there is diversity in market liquidity across the 28 markets, there were also periods of common behavior. For example, the August 2011 downgrade of U.S. Treasury debt by Standard & Poor's coincided with ongoing fiscal weakness in several eurozone countries and the start of the Occupy Wall Street movement to produce a sharp, but ultimately

[23]The MRHMC model provides a useful robustness check. In general, the MRHMC fits were less stable on our data sample, and we focus on the HMC results to present a simpler, unified approach in the subsequent hierarchical analysis (see Section 4.2 below).

transient, spike in the probability of the low-liquidity (high price-impact) state. Similarly, the liquidity crisis after the failure of Lehman Brothers is plainly visible as the deep and more persistent spike in September 2008, preceded by a series of pronounced foreshocks over the course of the year.

Figure 5 condenses the trivariate time series of Figure 4 into univariate daily color codes, where each of the 28 horizontal ribbons represents one market over time. Within each ribbon, each trading day appears as a thin vertical sliver. The color of each daily sliver mixes the three primary colors of Figure 4 as a linear combination of red/yellow/blue color vectors, weighted by their respective state probabilities on each day (black indicates a missing value). Figure 5 groups all 28 markets by asset class, and covers the full sample period. On any given day, one state, and therefore one color, tends to dominate, but there are occasional exceptions (i.e., pink, green, or orange days in Figure 5). The ribbon charts illustrate that the equity markets and the VIX® index responded strongly and immediately to the funding market distress in August 2007, but WTI futures did not. Throughout the 2007-09 crisis window, VIX® liquidity was more persistently stressed compared to the other asset classes. Corporate bond liquidity for the speculative-grade bonds took longer to recover from the elevated illiquidity levels of the crisis episode. Consistent with the increased uncertainty about the financial sector equities liquidity overall remained depressed from late 2007 through 2008 and most of 2009. Two other key insights from the visual inspection are that the liquidity implications of the Lehman Brothers failure were felt broadly for an extended period and that hints of illiquidity foreshocks existed in most markets, including most equities, WTI futures (perhaps reflecting the business cycle), and certain VIX® maturities. These patterns may ultimately help in crafting liquidity forecasts.

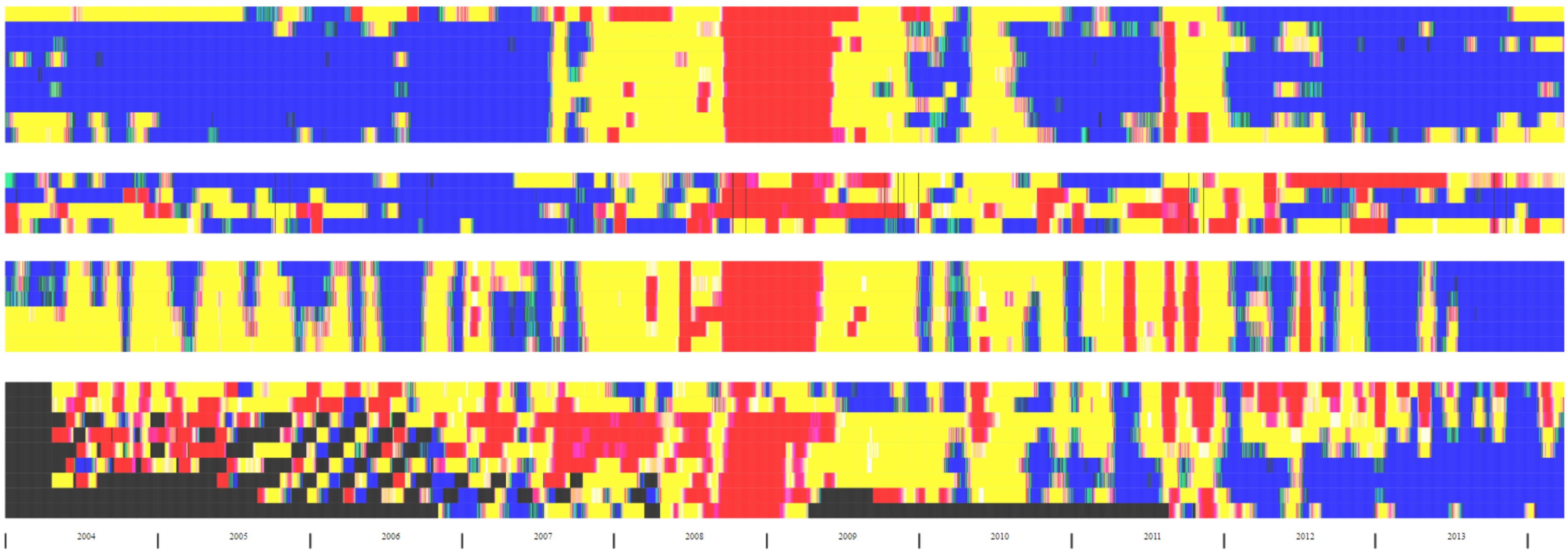


Figure 5: Daily Price-Impact Probabilities across all 28 Markets,
Top to bottom: Equities (SIC 0-8), Bonds (all grades), WTI futures, VIX® futures

Sources: Center for Research in Securities Prices, Bloomberg L.P., Mergent Inc., Wharton Research Data Services, Financial Industry Regulatory Authority, authors' analysis

An established alternative to our approach for modeling the price impacts across multiple markets is a vector auto-regressive (VAR) model, where the vector of price impacts across the various markets being considered are assumed to have an auto-correlation structure, meaning that lagged values of price impacts from one market potentially drive the current price impacts observed in some or all of the markets. Preliminary analysis of the equity markets, for example, revealed that the price impacts across these markets are highly correlated. Although the first factor in a principal components analysis "only" explains about 70 percent of the variation, Figure 6 shows that this factor captures the critical parts of the liquidity dynamics with respect to the financial crisis.

Although there is a high level of multicollinearity, the primary signal in the data is strong enough that each of the markets, conditional on its own history, is essentially independent. We find that the VAR model essentially reduces to a collection of independent autoregressive models where only the lagged values (up to at most two lags) for each market are statistically significant.[24] In addition to exploring own lags, we considered an extension of the VAR framework that included both the lagged values of the values for each market and the 11 systemwide market summaries, with various levels of lagging. In all cases, these 11 systemwide summaries were not statistically significant. In our view, while a VAR may offer some predictive power, it will identify and predict significant changes to market liquidity no more than two days after those changes begin to appear in the market. In addition, a VAR does not take advantage of the information included in the 11 systemwide market summaries, because a linear relationship does not exist. We conclude that the VAR framework is not the best tool for policy makers. If a linear model is to be used, a more parsimonious approach would be to model the primary latent factor, using a Bayesian factor model, and allow the transition dynamics of that latent factor to depend

[24]This was true even when we used the stochastic-search variable selection method proposed by George and McCulloch [1993], which achieves parsimony by removing variables with no relationships. See Koop and Korobilis [2010] for a general discussion of the VAR methods used.

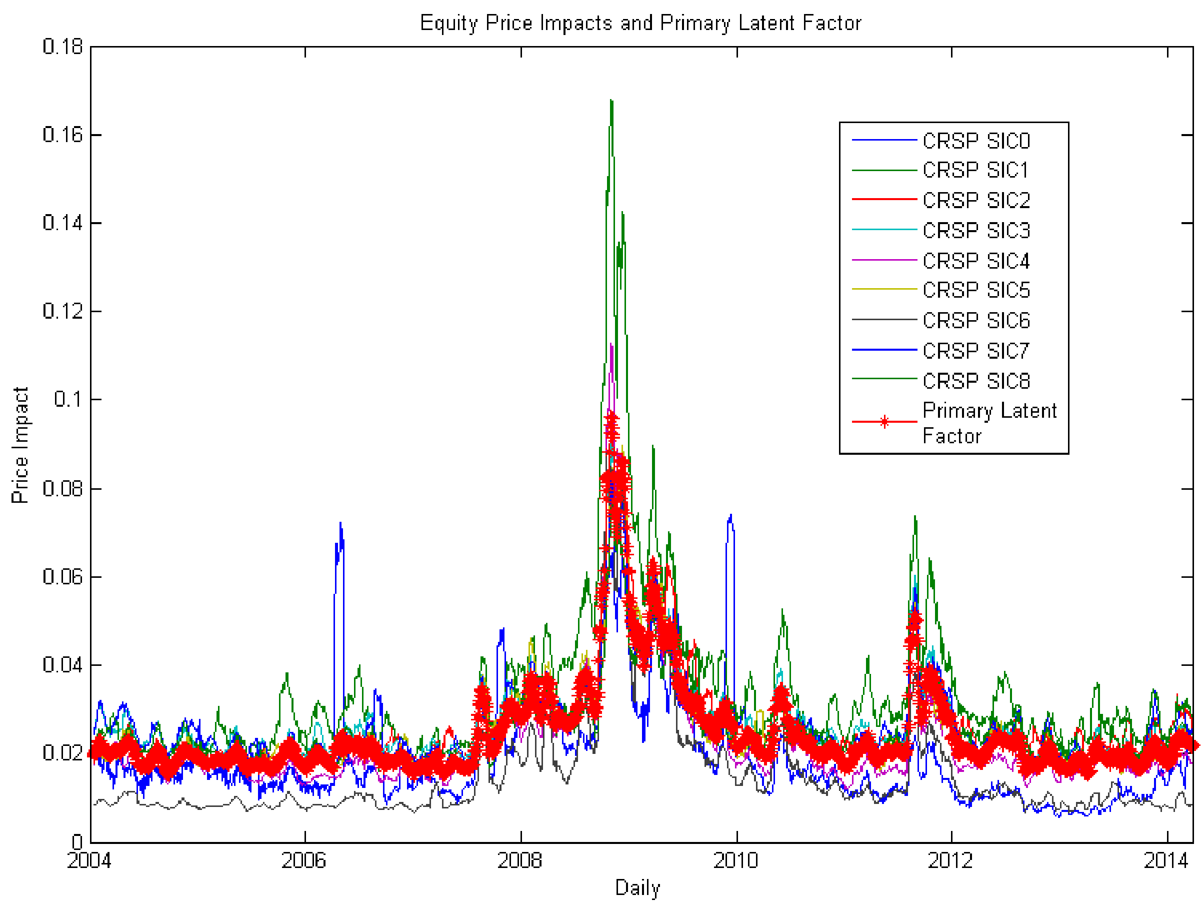


Figure 6: Price Impacts and Primary Liquidity Factor
Sources: Center for Research in Securities Prices, Wharton Research Data Services, authors' analysis

upon the summary series used in the multivariate, hidden Markov model.

4.2 Explaining Liquidity Regimes

There are strong relationships between changes in the level of liquidity and a number of summary series. Although the visual exploration of these relationships is helpful, the formal hierarchical model allows us to determine whether these relationships are statistically significant, particularly in the presence of other competing summary series.

We test the ability of the summary series to recover the liquidity dynamics across markets in two ways. First, we calculate a simple hit rate, which is the proportion of the time that the probit model, based solely on the summary series, accurately predicts the state identified by each of the underlying univariate models. That is, we count the proportion of time that we accurately predict the state of D_{it} for each i and t using the current estimate of β, Σ and the summary series data x_{it}. The naive hit rate, assuming random guessing, is 33 percent; in contrast, the posterior average of probit model's hit rate was 54 percent, indicating that the summary series are explaining a significant portion of the liquidity dynamics. Second, we compare the predictions of the liquidity state identified by the probit model to the liquidity state from the underlying MCMC data. As an illustration of the model fit, Figure 7 compares the fitted daily probit prediction (red) to the underlying HMC state probabilities (blue). In this example, we fit the model against the equities subsample, so the blue line in each panel of Figure 7 represents the average probability, across all nine SIC portfolios, of being in the relevant state. The predicted probabilities closely track the average probabilities, confirming the ability of the summary series to explain the liquidity dynamics.[25]

[25] As a robustness check, we also estimated the probit model on various subsets of the four asset classes in our sample. The hit rate is higher when narrowing the sample to equities only, or equities plus corporate bonds, suggesting that the broader panel may introduce some cross-sectional variation that is not fully captured by the probit. In both cases, the signs and magnitudes on the posterior means from the probit estimation are broadly similar. Similarly, the precision of the fit of the average probability from the four

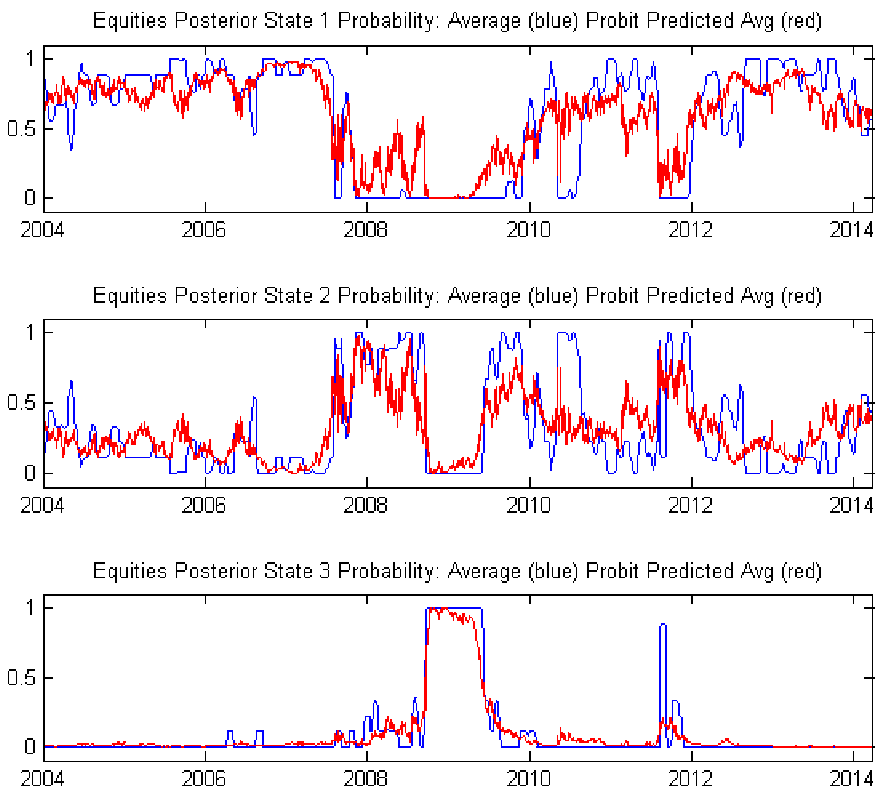


Figure 7: Average State Probabilities versus Probit Predicted Probabilities (Equities)
Sources: Center for Research in Securities Prices, Wharton Research Data Services, Bloomberg L.P, authors' analysis

Table 4 presents the probit model estimates. We standardized the inputs to the probit model by mean-centering each summary series and dividing it by its standard deviation, to facilitate direct comparisons of the magnitude of the parameter estimates. We force the latent value for state 1 to always be zero to address the additive identification restriction. Parameter estimates for states 2 and 3 are the difference between the unrestricted parameters of each of these states relative to state 1. The negative intercepts indicate that state 1, the low price-impact or high-liquidity state, is the most prevalent state when the associated summary variable is positive, and the fact that the intercept for state 3 is more negative than for state 2 indicates that state 3 (the low-liquidity state) is the least likely state when the associated summary variable is positive.

Table 4: Posterior Parameter Estimates Probit Portion of Hierarchical Model

	Posterior Mean		**Posterior StDev**		**t-Stat**	
Summary Variable	**State 2**	**State 3**	**State 2**	**State 3**	**State 2**	**State 3**
Intercept	-0.34**	-0.79**	0.01	0.01	-36.59	-94.01
VIX®	0.17**	0.18**	0.01	0.01	13.82	12.97
U.S. 5y Breakeven Inflation	0.13**	0.07**	0.01	0.01	6.40	3.94
WTI	0.15**	-0.10**	0.01	0.01	13.40	-21.32
3-month Repo Rate	0.47**	-0.52**	0.02	0.01	24.94	-56.07
TED Spread	0.35**	-0.04**	0.02	0.01	14.65	-4.47
Yield Spread (10y–2y)	0.25**	-0.46**	0.01	0.01	27.11	-58.25
S&P 500 Price-Book Ratio	0.05**	-0.08**	0.01	0.01	3.17	-5.76
Dow Jones Real Estate Index	-0.53**	0.07**	0.01	0.01	-76.23	24.84
Moody's Baa Bond Index	-0.30**	0.28**	0.01	0.01	-20.51	54.19
LIBOR–OIS Spread	-0.40**	0.11**	0.02	0.01	-11.93	9.19
U.S. Dollar Index	-0.32**	-0.19**	0.02	0.01	-45.05	-33.57

** Significant at a 99 percent confidence level
Sources: Center for Research in Securities Prices, Bloomberg L.P,
Wharton Research Data Services, authors' analysis

All of the summary series are statistically significant. Within these results, there are some interesting patterns to note. First, there is a natural grouping among the summary

bond portfolios, but somewhat less precisely. Nonetheless, the inclusion of the price-impact data for the bonds improves the predictive performance of the model.

series with regards to the pattern of the signs for the state 2 and 3 parameter estimates. As might be expected, VIX® has a positive-positive pattern indicating that higher levels of VIX® are associated with a higher probability of entering states with low liquidity. U.S. Breakeven Inflation exhibits a similar pattern, albeit with slightly smaller coefficients. Another group of five summary series (WTI, 3-month Repo Rate, S&P 500 Price-Book Ratio, TED Spread, and Yield Spread) exhibit a positive-negative pattern, indicating that elevated levels of these summary series are associated with a high probability of being in the middle liquidity state. This pattern may seem counter-intuitive, but it partly reflects the fact that the model does not anticipate the responses by central banks and market participants in the wake of the 2008 crisis. In other words, the model predicts that extreme illiquidity conditions (state 3) will persist post-crisis longer than they actually do, because the market and policy response was a significant departure from pre-crisis behavior. The next set of summary series (Dow Jones U.S. Real Estate Index, Moody's Baa Corporate Bond Index, and LIBOR-OIS spread) exhibit a negative-positive pattern, reflecting persistently high levels during episodes of low liquidity which then bounce back sufficiently during episodes of intermediate liquidity to have the extremes associated with the crisis. Finally, the U.S. Dollar Index shows a significant negative-negative pattern, indicating that higher levels of the dollar are associated with higher probability of entering a high-liquidity state. This is consistent with a flight to quality, in which capital flows into the United States during episodes of stress, simultaneously pushing up the value of the dollar and flooding the domestic market with liquidity.

To assist in our understanding of these parameter estimates, we can compare the time-series plot for individual summary series versus the probit-predicted probability for each state. As an illustration of the model, Figure 8 compares the TED spread, to the probit predictions of the three states (fitted again using the subsample of nine equities portfolios, as in Figure 7). The TED spread (blue line, repeated in all three panels) remains low until

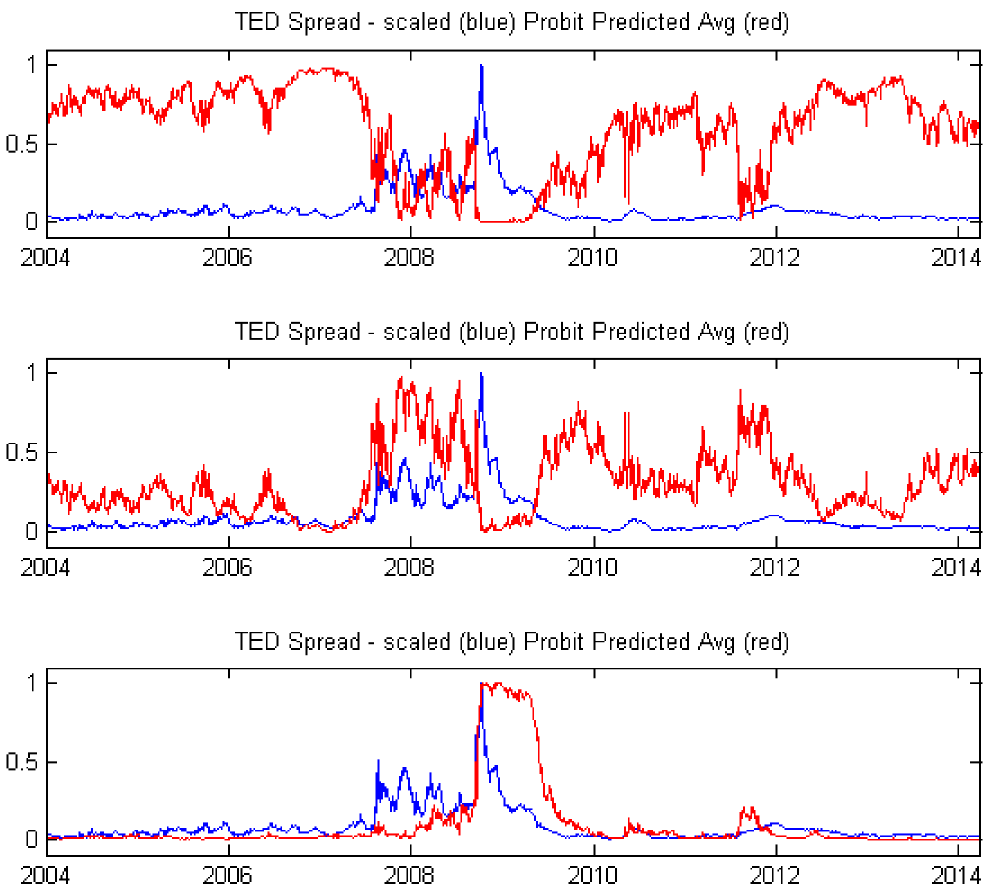


Figure 8: TED Spread versus Probit Probabilities
Sources: Center for Research in Securities Prices, Wharton Research Data Services, Bloomberg L.P., authors' analysis

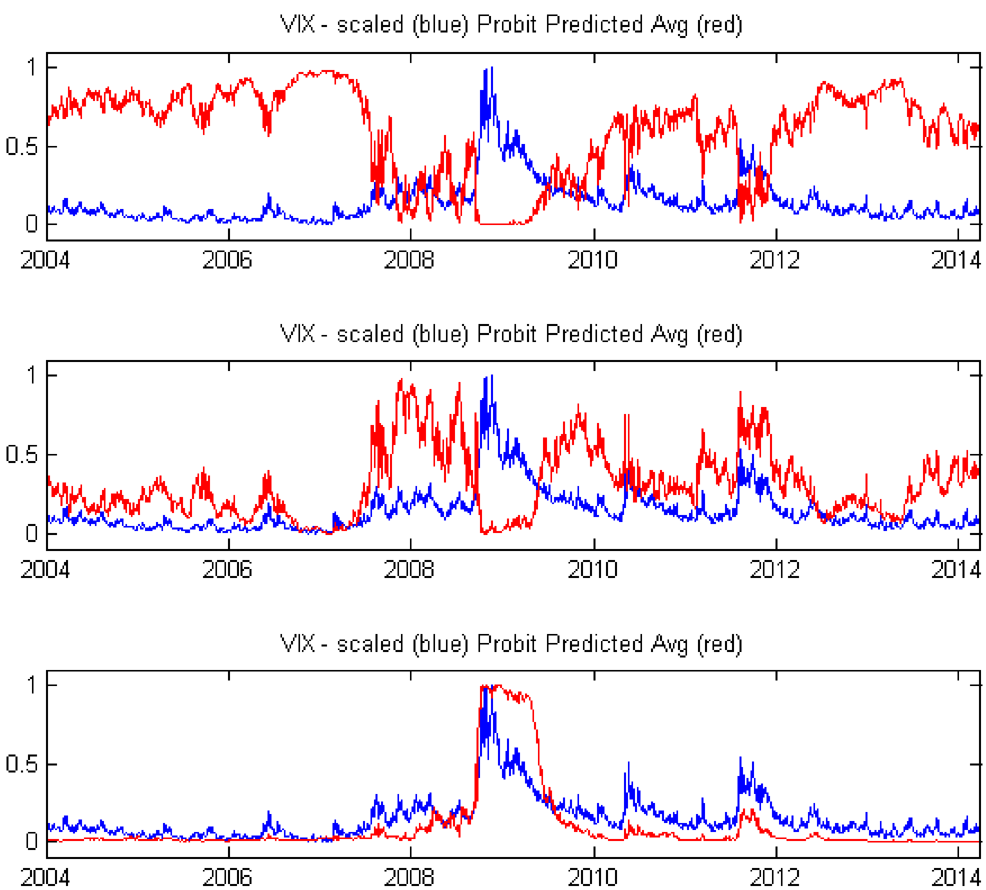


Figure 9: VIX® versus Probit Probabilities
Sources: Center for Research in Securities Prices, Wharton Research Data Services, Bloomberg L.P., authors' analysis

mid-2007 and returns to persistent low levels in 2010. The early episode corresponds to consistently high probabilities for high liquidity (state 1, the red line in the top panel), consistent with the TED spread's role as a bellwether for funding liquidity. Between August 2007 and September 2008, as the TED spread begins to widen, the probability of state 2 jumps (middle panel), supporting the positive coefficient in Table 4. After September 2008, the TED spread recedes relatively quickly from its peak, compared with the probability of being in state 3 (bottom panel), which remains elevated for the next year. As noted above, this deviation between the actual TED spread and the probit-predicted probability of extreme illiquidity is consistent with the negative coefficient on state 3 in Table 4. In contrast, the VIX® index has positive coefficients on both states 2 and 3 in Table 4. In contrast, Figure 9 shows that the VIX® is more persistently high after the 2008 shock, consistent with the positive coefficient on state 3 in Table 4. Moreover, the VIX® remains moderately elevated for much of the post-crisis period after 2009, when the probit-predicted probability for state 2 is also raised. The hierarchical model clearly provides insights into the relationship between the systemwide summary variables and the detailed market liquidity dynamics. The hierarchical model offers a valuable tool for understanding the drivers of liquidity across a broad range of markets.

5 Predicting Liquidity Regimes

Because of the difficulties providing policy makers with a parsimonious representation based a VAR or latent factor model, we choose to explore the predictive power of our proposed multivariate model and leave comparison of the relative predictive performance of these competing methods as a point for future research. The relatively high hit rate of the multivariate model, based on the equity portfolios, suggests that it may have reasonable predictive power, and in fact the high hit rate, combined with the high temporal correlation in the summary series offers what appears to be a potentially powerful tool for predicting future periods of high price impact, which correspond to times of severe financial stress.

There may be some concern that using the current value of the summary series to explain the current liquidity state could result in a model where any relationships that were found are either endogenous or driven by common unobserved factors. In exploring this issue, we found that adding lagged values of the summary series (beyond the current summary series) provides almost no additional benefit in recovering the average liquidity state. With this in mind, we have rerun the hierarchical analysis excluding current values of the summary series, and using only lagged values instead. Significantly, using information lagged by up to 15 days provides hit rates that are essentially the same as hit rates using concurrent information. The persistence in hit rates at increasing lags gives us reason to consider the possibility that the probit model may have forecasting power.[26]

To illustrate the potential predictive power of the multivariate hidden Markov model during a period of financial stress, we calibrated the model by lagging the summary series at four different levels, using equity, bond, and futures data from March 1, 2004 through June 29, 2007: (i) the current values (Lag 0); (ii) one day old (Lag 1); (iii) five days old (Lag 5); and (iv) 15 days old (Lag 15). We then used the parameters estimated from each analysis and the prevailing summary series going forward to predict the probability of being in each of the different liquidity regimes. For example, we input the July 1, 2007 summary data with the Lag 0 analysis and predict the probabilities of being in the different liquidity regimes for July 1, 2007. On the other hand, using the same summary data with the Lag 15 model we were predicting the probability of being in the different liquidity regime for July 16, 2007. As indicated by the dashed lines (which represent the predicted probabilities) in Figures 10 and 11, all the lagged models predict a jump to the high price impact or low

[26] As a robustness check, we also examined the persistence of hit rates at the various lags for various subsamples of the markets. The results were very similar, with hit rates remaining essentially unchanged as the lag increased.

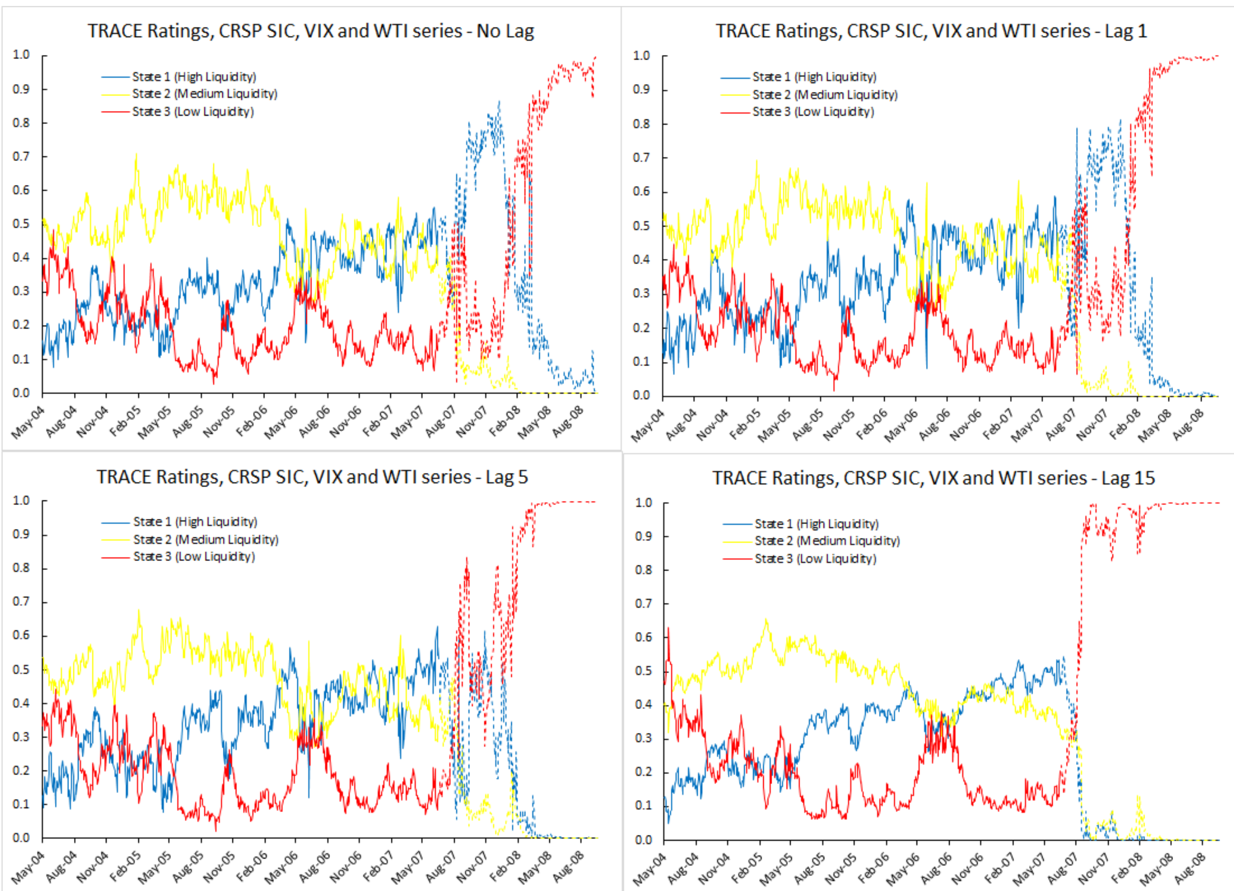


Figure 10: Predicted Liquidity States at Four Lags
Sources: Center for Research in Securities Prices, Wharton Research Data Services, Bloomberg L.P., Mergent, Inc., authors' analysis

liquidity regime well in advance of the crisis of 2008. It is noteworthy that the Lag 5 and Lag 15 predictions identify the period of stress sooner than the Lag 1 results. This happens in part because they skip forward and in part because they are more sensitive to extreme jumps — perhaps because there is some endogeneity in the current summary series data and the current liquidity states that is eliminated by lagging the summary series data at least five days.

Interestingly, when we consider the bond data alone in the HMC model, we find lower hit rates as well as a reduced ability to recover the average probabilities when compared to using equity data alone. We are also unable to recover the average probabilities as well as when using the equity data alone. However, when we analyze both the bond and the equity data together in the HMC model, there is a marked improvement in the model's performance — especially in its predictive ability. This suggests that the bond data contain information which differs from and supplements the information contained in the equity data.

Clearly these predictive results suggest that we explore extensions to this predictive approach and contrast it with alternative predictive methods. Part of our extension will include overcoming data consistency issues with price impact measurements from markets beyond the equity markets. One approach would be to create a Bayesian Factor model, where the missing price impacts are treated as unobserved (latent) values that can be estimated in a manner that is consistent with the factor structure that is uncovered. This Bayesian factor model, can then be integrated into an extended version of the multivariate hidden Markov model, where the a finite mixture structure is used to group markets based on both their loadings on the underlying factors and the dynamics of their hidden Markov chains. Clearly it would be of interest to compare this extended model with the predictive performance of a version of the Bayesian factor model where the dynamics of the factors are driven by the summary series that are used with the multivariate hidden Markov model.

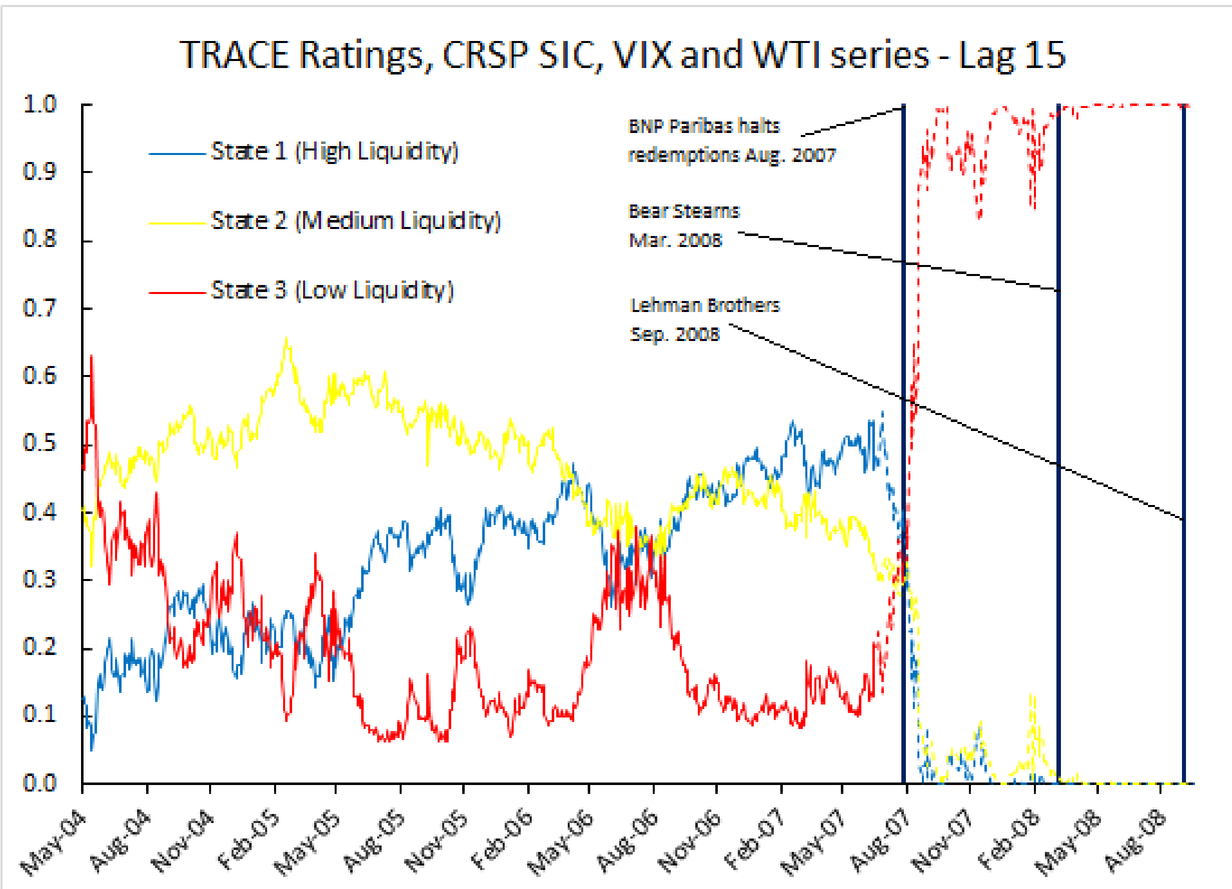


Figure 11: Predicted Liquidity States at Lag 15 using both Equity and Bond price impacts.
Sources: Financial Industry Regulatory Authority, Center for Research in Securities Prices, Wharton Research Data Services, Bloomberg L.P., authors' analysis

Finally, to explore the practical value of this approach to policy makers, we need to explore the performance of this modeling for a range of different periods of extreme liquidity shocks both across time and across different geographic areas. We anticipate that this exploration will lead us to conclude that the impact of different summary series on predicting liquidity regimes change over time, ultimately leading us to develop a comprehensive model that has a dynamic component, with regards to the parameters in the multinomial probit portion of the multivariate hidden Markov model.

6 Conclusion

Liquidity is an elusive, yet essential component of the modern financial system. It is elusive because conceptually it is hard to define, and empirically it is hard to measure and predict. We attribute the challenges in liquidity measurement to three fundamental aspects of the phenomenon. Liquidity is latent, in the sense that the episodes of illiquidity we seek to understand are rare, and often emerge with little apparent warning. Liquidity is nonlinear, in the sense that price impact does not respond proportionately to additional order flow, making it difficult to extrapolate from ordinary markets to the behavior of those markets under stress. Liquidity is endogenous, in the sense that it often emerges as a positive externality in very active markets, making those busy venues attractive to others who seek the assurance that counterparties will be available when needed.

We address the challenges of latency, nonlinearity and endogeneity statistically with a Bayesian estimation of a hidden Markov chain individually for 28 separate time series covering the CRSP and TRACE universes of U.S. equities and corporate bonds, plus multiple expiries of two key futures contracts, the VIX® volatility contract and the WTI oil contract. Three latent states (high, medium, and low price impact) are adequate to capture the observed liquidity structure of all 28 univariate series.

We look for cross-sectional structure in liquidity episodes by estimating a hierarchical Bayesian model, and testing the ability of several systemwide market summary time series to recover the estimated aggregate liquidity dynamics. This exercise also permits an attribution of those estimated aggregate dynamics to meaningful economic interpretations.

We also explore the predictive power of this model, using lagged values of the summary series, and find that the model offers a possible predictive tool for identifying future jumps in market liquidity as far out as 15 days in advance. Our results at this stage are preliminary, but also promising. In addition to testing for robustness and sensitivity, we see several immediate avenues for future research, including expanding the cross section of asset markets in the scope of analysis, comparing in more detail the liquidity behavior of wholesale funding markets, and experimenting with alternative portfolio formation rules.

References

Tobias Adrian and Hyun Song Shin. Liquidity and leverage. *Journal of Financial Intermediation*, 19(3):418–437, July 2010. URL http://www.sciencedirect.com/science/article/pii/S1042957308000764.

James H Albert and Siddhartha Chib. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679, 1993. doi: 10.1080/01621459.1993.10476321. URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1993.10476321.

Franklin Allen and Douglas Gale. *Understanding Financial Crises*. Oxford University Press, 2009. URL https://global.oup.com/academic/product/understanding-financial-crises-9780199251421?cc=us&lang=en&.

Y. Amihud. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5:31–56, 2002. URL https://umdrive.memphis.edu/pjain/Microstructure/Papers/Illiquidity%20and%20stock%20returns_Amihud.pdf.

Yakov Amihud and Haim Mendelson. Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223–249, 1986. doi: 10.1016/0304-405X(86)90065-6. URL http://www.sciencedirect.com/science/article/pii/0304405X86900656.

Yakov Amihud, Haim Mendelson, and Lasse Heje Pedersen. *Market Liquidity: Asset Pricing, Risk, and Crises*. Cambridge University Press, 2013. URL http://ebooks.cambridge.org/ebook.jsf?bid=CBO9780511844393.

Torben G. Andersen, Oleg Bondarenko, Albert S. Kyle, and Anna A. Obizhaeva. Intraday trading invariance in the e-mini s&p 500 futures market. Technical report, U. of Maryland, March 2016. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2693810.

Niki Anderson, Lewis Webber, Joseph Noss, Daniel Beale, and Liam Crowley-Reidy. The resilience of financial market liquidity. Financial Stability Paper 34, Bank of England, October 2015. URL http://www.bankofengland.co.uk/financialstability/Documents/fpc/fspapers/fs_paper34.pdf.

Michael J. Barclay, Terrence Hendershott, and Kenneth Kotz. Automation versus intermediation: Evidence from treasuries going off the run. *Journal of Finance*, 61(5):2395–2414, October 2006. ISSN 00221082. URL `http://www.jstor.org/stable/3874714`.

Leonard E. Baum, Ted Petrie, George Soules, and Norman Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *Annals of Mathematical Statistics*, 41(1):164–171, 1970. URL `http://projecteuclid.org/euclid.aoms/1177697196`.

Hendrik Bessembinder, William Maxwell, and Kumar Venkataraman. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics*, 82(2):251–288, November 2006. URL `http://www.sciencedirect.com/science/article/pii/S0304405X06000699`.

Hendrik Bessembinder, Marios Panayides, and Kumar Venkataraman. Hidden liquidity: an analysis of order exposure strategies in electronic stock markets. *Journal of Financial Economics*, 94(3):361–383, December 2009. URL `http://www.sciencedirect.com/science/article/pii/S0304405X09001494#`.

Tim Bollerslev. Glossary to ARCH (GARCH). In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, Advanced Texts in Econometrics, chapter 8, pages 137–163. Oxford University Press, 2010. URL `https://global.oup.com/academic/product/volatility-and-time-series-econometrics-9780199549498?cc=us&lang=en&`.

Richard Bookstaber, Michael D. Foley, and Brian F. Tivnan. Market liquidity and heterogeneity in the investor decision cycle. OFR Working Paper 15-02, Office of Financial Research, February 2015.

Jean-Philippe Bouchaud, J. Doyne Farmer, and Fabrizio Lillo. How markets slowly digest changes in supply and demand. In T. Hens and K. Schenk-Hoppe, editors, *Handbook of Financial Markets: Dynamics and Evolution*, page 57156. Academic Press, 2008. URL `http://arxiv.org/pdf/0809.0822.pdf`.

Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. CRC press, 2011. URL `http://www.crcpress.com/product/isbn/9781420079425`.

M. Brunnermeier and L.H. Pedersen. Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6):2201–2238, June 2009. URL `http://rfs.oxfordjournals.org/content/22/6/2201.abstract`.

Markus K. Brunnermeier. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives*, 23(1):77–100, Winter 2009. URL `http://ideas.repec.org/a/aea/jecper/v23y2009i1p77-100.html`.

Olivier Cappé, Eric Moulines, and Tobias Rydén. *Inference in Hidden Markov Models*. Springer Verlag, 2005. URL http://link.springer.com/book/10.1007/0-387-28982-8.

Chris K. Carter and Robert Kohn. On gibbs sampling for state space models. *Biometrika*, 81(3):541–553, September 1994. URL http://biomet.oxfordjournals.org/content/81/3/541.short.

Sally Chen, Philip Liu, Andrea Maechler, Chris Marsh, Sergejs Saksonovs, and Hyun Song Shin. Exploring the dynamics of global liquidity. Technical report, IMF, 2012. URL http://www.imf.org/external/pubs/ft/wp/2012/wp12246.pdf.

Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Commonality in liquidity. *Journal of Financial Economics*, 56(1):3–28, 2000. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.139.1722&rep=rep1&type=pdf.

Tri Vi Dang, Gary Gorton, and Bengt Holmström. Financial crises and the optimality of debt for liquidity provision. Technical report, Yale U., May 2010. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.183.7960&rep=rep1&type=pdf.

Piet De Jong and Neil Shephard. The simulation smoother for time series models. *Biometrika*, 82(2):339–350, June 1995. URL http://biomet.oxfordjournals.org/content/82/2/339.short.

Hans Degryse, Frank De Jong, and Vincent Van Kervel. The impact of dark trading and visible fragmentation on market quality. *Review of Finance*, page rfu027, 2014. URL http://rof.oxfordjournals.org/content/early/2014/06/13/rof.rfu027.short.

J. Bradford DeLong, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance*, 45(2):379–395, June 1990. ISSN 00221082. URL http://www.jstor.org/stable/2328662.

Jens Dick-Nielsen. Liquidity biases in trace. *Journal of Fixed Income*, 19(2):43–55, Fall 2009. doi: DOI: 10.3905/jfi.2009.19.2.043. URL http://www.iijournals.com/doi/abs/10.3905/jfi.2009.19.2.043.

Feike C. Drost and Theo E. Nijman. Temporal aggregation of garch processes. *Econometrica*, 61(4):909–927, July 1993. URL http://www.jstor.org/stable/2951767.

Darrell Duffie. Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance*, 65(4):1237–1267, 2010. URL http://www.darrellduffie.com/uploads/pubs/DuffieAFAPresidentialAddress2010.pdf.

David Easley and Maureen O'Hara. Time and the process of security price adjustment. *Journal of Finance*, 47(2):577–605, June 1992. ISSN 00221082. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1992.tb04402.x/abstract.

Nicholas Economides. The economics of networks. *International Journal of Industrial Organization*, 14(6):673–699, October 1996. URL http://www.sciencedirect.com/science/article/pii/0167718796010156.

Robert F. Engle. The econometrics of ultra-high-frequency data. *Econometrica*, 68(1):1–22, January 2000. ISSN 00129682. URL http://www.jstor.org/stable/2999473.

Robert F Engle and Joe Lange. Predicting vnet: A model of the dynamics of market depth. *Journal of Financial Markets*, 4(2):113–142, April 2001. URL http://www.sciencedirect.com/science/article/pii/S1386418100000197.

FINRA. Historic trace data: Enhanced historic time and sales trade record file layout. Technical report, FINRA, 2009. URL http://www.finra.org/web/groups/industry/@ip/@comp/@mt/documents/appsupportdocs/p120989.pdf.

Sean Foley, Katya Malinova, and Andreas Park. Dark trading on public exchanges. Technical report, University of Toronto, February 2013. URL http://ssrn.com/abstract=2182839.

Thierry Foucault, Marco Pagano, and Ailsa Röell. *Market Liquidity: Theory, Evidence, and Policy*. Oxford University Press, 2013. URL http://ukcatalogue.oup.com/product/9780199936243.do.

Xavier Gabaix, Parameswaran Gopikrishnan, Vasiliki Plerou, and H Eugene Stanley. Institutional investors and stock market volatility. *Quarterly Journal of Economics*, 121(2):461–504, 2006. URL http://qje.oxfordjournals.org/content/121/2/461.abstract.

Alexandros Gabrielsen, Massimiliano Marzo, and Paolo Zagaglia. Measuring market liquidity: An introductory survey. Working Paper DSE 802, U. di Bologna, 2011. URL http://arxiv.org/pdf/1112.6169.pdf.

Deborah Gefang, Gary Koop, and Simon M. Potter. Understanding liquidity and credit risks in the financial crisis. *Journal of Empirical Finance*, 18(5):903–914, December 2011. URL `http://www.sciencedirect.com/science/article/pii/S0927539811000508`.

Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, 3rd edition, 2013. URL `http://www.amazon.com/dp/1439840954/`.

Edward I George and Robert E McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993. doi: 10.1080/01621459.1993.10476353. URL `http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1993.10476353#.VMorSfldVCY`.

Joel Hasbrouck. Intraday price formation in us equity index markets. *Journal of Finance*, 58(6):2375–2400, December 2003. URL `http://onlinelibrary.wiley.com/doi/10.1046/j.1540-6261.2003.00609.x/abstract`.

Joel Hasbrouck. *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*. Oxford University Press, 2007. URL `https://global.oup.com/academic/product/empirical-market-microstructure-9780195301649`.

Joel Hasbrouck and Gideon Saar. Technology and liquidity provision: The blurring of traditional definitions. *Journal of Financial Markets*, 12(2):143–172, 2009. ISSN 1386-4181. doi: http://dx.doi.org/10.1016/j.finmar.2008.06.002. URL `http://www.sciencedirect.com/science/article/pii/S1386418108000220`.

Joel Hasbrouck and Duane J Seppi. Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59(3):383–411, 2001. URL `http://www.sciencedirect.com/science/article/pii/S0304405X0000091X`.

Craig W. Holden, Stacey E. Jacobsen, and Avanidhar Subrahmanyam. The empirical analysis of liquidity. *Foundations and Trends in Finance*, forthcoming, 2014. doi: 10.2139/ssrn.2402215. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402215`.

Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91, March 1993. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1993.tb04702.x/full`.

Narasimhan Jegadeesh and Sheridan Titman. Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance*, 56(2):699–720, April 2001. URL `http://onlinelibrary.wiley.com/doi/10.1111/0022-1082.00342/abstract`.

Rudolf E. Kálmán. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, March 1960. doi: 10.1115/1.3662552. URL `http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1430402`.

G. Andrew Karolyi, Kuan-Hui Lee, and Mathijs A. van Dijk. Understanding commonality in liquidity around the world. *Journal of Financial Economics*, 105:82–112, 2012. URL `http://www.sciencedirect.com/science/article/pii/S0304405X11002844`.

Ira G. Kawaller, Paul D. Koch, and Timothy W. Koch. The temporal price relationship between s&p 500 futures and the s&p 500 index. *Journal of Finance*, 42(5):1309–1329, December 1987. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1987.tb04368.x/abstract`.

Gary Koop and Dimitris Korobilis. *Bayesian multivariate time series methods for empirical macroeconomics*. Now Publishers Inc., 2010. URL `https://books.google.com/books?id=JcVn9IJ9wSsC`.

A. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, November 1985. URL `https://www.econometricsociety.org/publications/econometrica/1985/11/01/continuous-auctions-and-insider-trading`.

Albert S. Kyle and Anna A. Obizhaeva. Large bets and stock market crashes. working paper, University of Maryland, May 2013. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2023776`.

Albert S. Kyle and Anna A. Obizhaeva. Market microstructure invariance: Empirical hypotheses. *Econometrica*, forthcoming 2016. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2722524`.

Sriketan Mahanti, Amrut Nashikkar, Marti Subrahmanyam, George Chacko, and Gaurav Mallik. Latent liquidity: A new measure of liquidity, with an application to corporate bonds. *Journal of Financial Economics*, 88(2):272–298, May 2008. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/j.jfineco.2007.02.006. URL `http://www.sciencedirect.com/science/article/pii/S0304405X08000172`.

Miriam Marra. Illiquidity commonality across equity and credit markets. working paper, University of Reading, April 2013. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2019512`.

Robert McCulloch and Peter E. Rossi. An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64(1):207–240, September-October 1994. URL `http://www.sciencedirect.com/science/article/pii/0304407694900647`.

Robert E. McCulloch, Nicholas G. Polson, and Peter E. Rossi. A bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics*, 99(1):173 – 193, 2000. ISSN 0304-4076. doi: http://dx.doi.org/10.1016/S0304-4076(00)00034-8. URL `http://www.sciencedirect.com/science/article/pii/S0304407600000348`.

Stephen Morris and Hyun Song Shin. Liquidity black holes. *Review of Finance*, 8(1):1–18, 2004. URL `http://cowles.econ.yale.edu/P/cp/p11a/p1114.pdf`.

Ulrich A. Müller, Michel M. Dacorogna, Richard B. Olsen, Olivier V. Pictet, Matthias Schwarz, and Claude Morgenegg. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking and Finance*, 14(6):1189–1208, 1990. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/0378-4266(90)90009-Q. URL `http://www.sciencedirect.com/science/article/pii/037842669090009Q`.

Mahendrarajah Nimalendran and Sugata Ray. Informational linkages between dark and lit trading venues. *Journal of Financial Markets*, 17:230–261, January 2014. ISSN 1386-4181. doi: 10.1016/j.finmar.2013.02.003. URL `http://www.sciencedirect.com/science/article/pii/S1386418113000128`.

NYU Volatility Institute (V-Lab). V-lab: Models. Technical report, New York University, 2014. URL `http://vlab.stern.nyu.edu/doc?topic=mdls`. Accessed 29 December 2014.

Marco Pagano. Trading volume and asset liquidity. *Quarterly Journal of Economics*, 104(2):255–274, May 1989. URL `http://qje.oxfordjournals.org/content/104/2/255.short`.

Christine A. Parlour and Duane J. Seppi. Limit order markets: A survey. In A.W.A. Boot and A.V. Thakor, editors, *Handbook of Financial Intermediation & Banking*, volume 5. Elsevier, 2008. URL `http://www.edegan.com/pdfs/Parlour%20Seppi%20(2008)%20-%20Limit%20Order%20Markets%20A%20Survey.pdf`.

Richard Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139, September 1984. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1984.tb03897.x/full`.

Hans R Stoll. Friction. *Journal of Finance*, 55(4):1479–1514, August 2000. URL http://onlinelibrary.wiley.com/doi/10.1111/0022-1082.00259/abstract.

Jean Tirole. Illiquidity and all its friends. *Journal of Economic Literature*, 49(2):287–325, 2011. URL http://idei.fr/doc/by/tirole/illiquidity-may12_11.pdf.

Bence Toth, Yves Lemperiere, Cyril Deremble, Joachim De Lataillade, Julien Kockelkoren, and J-P Bouchaud. Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2):021006/1–11, 2011. URL http://prx.aps.org/pdf/PRX/v1/i2/e021006.

Dimitri Vayanos and Jiang Wang. Market liquidity – theory and empirical evidence. NBER Working Paper 18251, National Bureau of Economic Research, July 2012. URL http://www.nber.org/papers/w18251.

Haoxiang Zhu. Do dark pools harm price discovery? *Review of Financial Studies*, 27(3):747–789, March 2013. doi: 10.1093/rfs/hht078. URL http://dspace.mit.edu/openaccess-disseminate/1721.1/88124.