Liquidity Risk After 20 Years

Ľuboš Pástor

Robert F. Stambaugh*

April 14, 2019

Abstract

The *Critical Finance Review* commissioned Li, Novy-Marx, and Velikov (2017) and Pontiff and Singla (2019) to replicate the results in Pástor and Stambaugh (2003). Both studies successfully replicate our market-wide liquidity measure and find similar estimates of the liquidity risk premium. In the sample period after our study, the liquidity risk premium estimates are even larger, and the liquidity measure displays sharp drops during the 2008 financial crisis. We respond to both replication studies and offer some related thoughts, such as when to use our traded versus non-traded liquidity factors and how to improve the precision of liquidity beta estimates.

---

*Pástor is at the University of Chicago Booth School of Business, 5807 S Woodlawn Ave, Chicago, IL 60637. Stambaugh is at the Wharton School of the University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA, 19104. Both authors are also at the NBER. Pástor is also at the CEPR and the National Bank of Slovakia. Email: lubos.pastor@chicagobooth.edu, stambaugh@wharton.upenn.edu. The views in this paper are the responsibility of the authors, not the institutions with which they are affiliated. For helpful comments, we thank Doug Diamond, Ralph Koijen, Stefan Nagel, Robert Novy-Marx, Jeff Pontiff, Jeff Russell, Rob Vishny, and the seminar audience at the University of Chicago.

1. Introduction

We began our work on liquidity risk 20 years ago, shortly after the 1998 collapse of the Long-Term Capital Management (LTCM) hedge fund. LTCM held multiple large trading positions that were exposed to the risk that market liquidity might deteriorate. One famous example was their long-short position in off-the-run versus on-the-run Treasury bonds. These positions performed well in normal times, but they suffered when liquidity dried up in the summer of 1998. The LTCM episode highlighted the facts that market liquidity fluctuates over time and that asset prices respond to these fluctuations.

Twenty years ago, the role of liquidity in asset pricing was viewed largely in *static* terms, based on differences in liquidity across assets. As in Amihud and Mendelson (1986), less liquid assets were understood to offer higher expected returns to compensate for their lower liquidity. Inspired by the LTCM episode, we took a *dynamic* view, focusing on time variation in liquidity. Our idea was that the price of an asset can depend not only on the asset's liquidity but also on the sensitivity of the asset's return to pervasive liquidity shocks.

Our published study, Pástor and Stambaugh (2003), makes three primary contributions. First, it advances the hypothesis that aggregate liquidity is a state variable relevant for asset pricing. Motivated by the evidence of commonality in liquidity (Chordia, Roll, and Subrahmanyam, 2000), we hypothesize that liquidity risk is systematic, or non-diversifiable, and thus potentially priced. Second, to test the above hypothesis, we construct a measure of aggregate stock market liquidity. From that measure, we derive a *liquidity factor* that captures innovations in aggregate liquidity. Finally, we find that the liquidity factor is priced, in that stocks with higher sensitivities to this factor—i.e., higher *liquidity betas*—have higher average returns. Liquidity betas are defined as slope coefficients $\beta_i^{\mathcal{L}}$ from the regression

$$r_{i,t} = \beta_i^0 + \beta_i^{\mathcal{L}} \mathcal{L}_t + \beta_i^{\mathrm{M}} \mathrm{MKT}_t + \beta_i^{\mathrm{S}} \mathrm{SMB}_t + \beta_i^{\mathrm{H}} \mathrm{HML}_t + \epsilon_{i,t}, \tag{1}$$

where $r_{i,t}$ denotes asset i's excess return, $\mathcal{L}_t$ is the liquidity factor, and MKT, SMB, and HML are the three factors of Fama and French (1993).

While our study focuses on equities, subsequent studies find liquidity risk to be priced in other asset classes, such as hedge funds (e.g., Sadka, 2010), private equity (e.g., Franzoni, Nowak, and Phalippou, 2012), emerging markets (e.g., Bekaert, Harvey, and Lundblad, 2007), and corporate bonds (e.g., Bongaerts, de Jong, and Driessen, 2017). Important theoretical insights into liquidity risk are provided by Acharya and Pedersen (2005), Brunnermeier and Pedersen (2009), and others. The literature on liquidity risk has grown tremendously, and we do not attempt to review it here.

A decade after LTCM, academic interest in liquidity risk was further stimulated by the 2008 financial crisis, during which markets experienced dramatic reductions in liquidity. The crisis highlighted the effects of liquidity shocks on not only asset prices but also real activity. When liquidity dries up in the market, other bad things tend to happen. Liquidity shocks can destroy value, for example, by creating runs on financial institutions or by impeding the financing of viable investment projects. The crisis clearly showed why investors may be willing to pay more for assets whose values better withstand market-wide liquidity shocks. Our liquidity factor picks up the key events of the crisis, as we show in Section 2.2. Because it happened five years after the publication of our paper, the crisis provides some out-of-sample validation for the methodology behind our liquidity factor.

We are grateful to the editor of the *Critical Finance Review*, Ivo Welch, for commissioning two studies to assess the extent to which our empirical results can be replicated. We are especially grateful to the authors of those studies, namely Hongtao Li, Robert Novy-Marx, Mihail Velikov, Jeffrey Pontiff, and Rohit Singla, for undertaking this task and executing it with the utmost competence and professionalism. We have learned a lot from their work.

In the rest of this paper, we first review key results of the replicators and respond to some of their claims. We then present a few thoughts on our liquidity factor, such as when to use its traded versus non-traded version, why we pay more attention to historical rather than predicted liquidity betas, and why our liquidity measure is more useful at the aggregate rather than the individual-stock level. Finally, we discuss the implications of the asymmetric nature of liquidity shocks. For example, we explain why liquidity betas are relatively hard to estimate and what can be done to improve their precision.

2. Replication Results

Both Li, Novy-Marx, and Velikov (2017) and Pontiff and Singla (2019) successfully replicate the construction of our liquidity factor over the time period used in our study (1962 through 1999). Li, Novy-Marx, and Velikov report a correlation of 98.9% between their liquidity series and ours. Pontiff and Singla report a correlation of one, up to five significant digits.

The above correlations pertain to our main liquidity factor, $\mathcal{L}_t$ from equation (1). Both studies also replicate the traded version of our factor, which we discuss in more detail in Section 3.1. Li, Novy-Marx, and Velikov (2017) report a 95% correlation between their traded liquidity factor and ours in 1968 through 1999. Both their factor and ours produce very similar alphas (see their Table 2). For example, both factors produce in-sample four-factor Carhart alphas that round to 0.37% per month and five-factor Fama-French alphas that round to 0.20% per month in 1968 through 1999. Pontiff and Singla (2019) also report similar results for the traded factor. Their in-sample estimate of the three-factor Fama-French alpha of the traded factor is 3.91% per year ($t = 2.00$; see their Table 1), which is close to the estimate of 4.15% ($t = 2.08$) reported in Table 8 of our 2003 study. Overall, both studies have managed to closely replicate our main in-sample results.

In general, a successful outcome of in-sample replication is not guaranteed even if the methodology used by the replication study is identical to that used in the original study. Even if the code is the same, the data need not be. Our 2003 study relies heavily on data from the Center for Research in Security Prices (CRSP). In an effort to keep its database as clean as possible, CRSP performs corrections to its historical data anytime it identifies, or becomes aware of, a data error. As a result, the CRSP dataset analyzed in a given year is not identical to the CRSP dataset covering the same period but downloaded years later. The two datasets are very similar, but the cumulative effect of CRSP updates over the interim period need not be negligible, especially if a long time has elapsed. A good example is provided in Table 1 of Pontiff and Singla. Their estimated in-sample alpha of 3.91% per year ($t = 2.00$), mentioned earlier, is identical to the alpha they obtain when they replace their replicated traded factor series with the updated series downloaded from our own websites. In other words, the difference between 3.91% and 4.15%, discussed in the previous paragraph, is not a result of a discrepancy between our methodology and that of Pontiff and Singla. Instead, it is driven by the corrections in the 1962–1999 CRSP dataset since we first used the data almost twenty years ago. Pontiff and Singla themselves write that "This minor difference is almost certainly attributable to year-to-year corrections to the CRSP data."

2.1. Out-of-Sample Performance

Our 2003 study estimates the liquidity risk premium based on the sample period of 1968 through 1999. If our estimate were a fluke, it would likely disappear out of sample. In contrast, both replication teams report that the magnitude of the liquidity risk premium has increased since 1999.

Li, Novy-Marx, and Velikov (2017) compare the in-sample (1968–1999) and the out-of-sample (2000–2015) performance of the traded liquidity factor, for both their replicated version and ours. They find that the out-of-sample performance is substantially stronger than the in-sample performance, for all three factor models they use. For example, the four-factor Carhart alpha of our factor increases from 0.37% per month ($t = 2.14$) in 1968–1999 to

0.68% per month ($t = 2.31$) in 2000–2015. Pontiff and Singla (2019) also report an improved post-sample performance of the traded factor, going through 2017 (see their Table 2). This evidence shows that had we written our 2003 study today, our main results would have been quite a bit stronger!

A possible reason why the evidence is stronger post-sample than in-sample is the 2008 financial crisis. This post-sample crisis produced wide fluctuations in liquidity, allowing more precise estimation of liquidity betas. In periods without liquidity crises, liquidity betas are estimated with less precision. We elaborate on this point in Section 4.

Given the strong post-sample performance of our traded liquidity factor, it is not surprising that the factor's full-sample performance in 1968 through 2015 is highly significant. In their Table 3, Li, Novy-Marx, and Velikov report that our factor's alphas with respect to the three-, four-, and five-factor models range from 0.44% to 0.46% per month, with t-statistics ranging from 2.85 to 3.02. The premium for liquidity risk seems alive and well.

Pontiff and Singla (2019) estimate not only the post-sample but also the pre-sample magnitude of the liquidity risk premium. Such analysis was not feasible when we wrote our 2003 study because CRSP added the pre-1962 daily data only in 2005. In the sample period of 1932 through 1967, Pontiff and Singla find no significant premium associated with liquidity risk: their traded factor has a premium of -2.72% per year, with the t-statistic of -1.36. As a result of this negative pre-sample performance, their estimate of the full-sample premium in 1932 through 2017 is positive but insignificant (see their Table 2).

Why is the pre-1962 evidence so different from the post-1962 evidence? Figure 1 of Pontiff and Singla (2019) shows that the pre-1962 fluctuations in liquidity are much smaller in magnitude than the more recent fluctuations. The largest pre-1962 fluctuations in liquidity are associated with the Great Depression, which seems plausible, but the scale of these fluctuations is substantially smaller than the scale of the post-1962 fluctuations.

One possible reason why the pre-1962 liquidity series is so much less volatile is a different stock universe. Whereas the post-1962 CRSP daily data include stocks from both the New York Stock Exchange (NYSE) and American Stock Exchange, the pre-1962 daily data includes only NYSE stocks. The larger average size of NYSE stocks could very well be responsible for the significantly lower pre-1962 volatility of the liquidity series. The source of the pre-1962 daily data is also different. The primary source is *The New York Times*, with the secondary source being *The Wall Street Journal*. The main source of post-1962 data is the Interactive Data Corporation. Given the different stock universes and data sources, the pre-1962 results may not be directly comparable to the post-1962 results. Additional analysis of the volatility differences in the pre- and post-1962 samples seems warranted.

2.2. The 2008 Financial Crisis

Our 2003 study came out five years before a major financial crisis, during which liquidity evaporated from many financial markets. The 2008 crisis provides a valuable out-of-sample perspective on the behavior of our liquidity factor. If the factor truly captures liquidity, it should have large negative realizations during the crisis. Luckily for the factor, it does.

The 2008 financial crisis is commonly considered to have begun in the summer of 2007 and ended in the spring of 2009. Between July 2007 and April 2009, our liquidity factor experienced negative innovations in 17 out of 22 months. Three of the five largest negative innovations during this period are associated with the defining moments of the crisis, namely, the demise of Lehman Brothers (September 2008), the fall of Bear Stearns (March 2008), and the quant crisis (August 2007). The two remaining largest drops in our liquidity measure occurred in December 2008 and January 2009, two of the worst months of the crisis. All five of these drops rank among the 13 biggest drops in our liquidity measure since 1962, and four of them are among the top 9. This evidence suggests that our liquidity factor, whose methodology was designed well before the 2008 crisis, does a good job capturing the most salient events of the crisis.

The large drops in liquidity are clearly visible in Figure 1, which plots the time series of our aggregate liquidity measure in August 1962 through December 2017. The drops in liquidity experienced during the 2008 crisis are not among the three largest drops in the full sample, which occurred in October 1987, September 1998, and November 1973. However, all three of those liquidity drops were short-lived. In contrast, during the 2008 crisis, liquidity diminished month after month, with devastating cumulative effects. For example, our liquidity measure experienced four large drops during the five-month period between September 2008 and January 2009; as a result, the cumulative drop in liquidity during this period was very large.

Our traded liquidity factor, which offers a long-short portfolio exposure to liquidity risk, also performed poorly during the financial crisis. A simple way to see this effect is in Figure 2 of Li, Novy-Marx, and Velikov (2017), which plots the cumulative performance of the traded factor in 1968 through 2015. The factor's performance has been strong overall, as indicated by the upward trend. Across the whole sample, the sharpest drop in the factor's performance occurs in the second half of 2008. Consider a strategy that invests $1 in a Treasury bill and takes a long-short position of $1 each in our traded factor, rebalanced monthly. The strategy's cumulative excess return over the six-month period (July through December 2008) is -34.3%, with negative returns in five of the six months (all except August). This is what one would expect from a strategy designed to capture the liquidity risk premium—it performs well on average, but it loses money in periods in which the risk materializes. These results provide additional out-of-sample support for our methodology.

2.3. Excluding Zero-Volume Days

To successfully replicate our liquidity factor, one must be aware of the fact that when we estimate equation (1) in our 2003 study, we exclude days with zero trading volume. That equation (1) takes the form of the regression

$$r^e_{i,d+1,t} = \theta_{i,t} + \phi_{i,t} r_{i,d,t} + \gamma_{i,t} \text{sign}(r^e_{i,d,t}) \cdot v_{i,d,t} + \epsilon_{i,d+1,t}, \quad d = 1, \ldots, D, \tag{2}$$

where $\gamma_{i,t}$ is the liquidity measure for stock i in month t and the remaining quantities are defined as follows:

- $r_{i,d,t}$: the return on stock i on day d in month t,
- $r^e_{i,d,t}$: $r_{i,d,t} - r_{m,d,t}$, where $r_{m,d,t}$ is the return on the CRSP value-weighted market return on day d in month t, and
- $v_{i,d,t}$: the dollar volume for stock i on day d in month t.

When we run this regression, we impose the screen $v_{i,d,t} > 0$, thereby excluding days with $v_{i,d,t} = 0$ from the analysis. Unfortunately, we inadvertently neglected to mention this important detail in our paper, as pointed out to us by Li, Novy-Marx, and Velikov (2017). If one replicates our analysis without imposing the positive-volume screen, the resulting series has only a modest correlation with our factor: 39.2% when computed over the period of August 1962 through December 1999, according to Li, Novy-Marx, and Velikov (2017), and 27% over the period of August 1962 through December 2017, according to Pontiff and Singla (2019). Our failure to mention the positive-volume screen may have caused some past replication attempts by other researchers to fail. We sincerely apologize for this omission. We are grateful to both replication teams for recognizing this oversight in a gracious manner.

Without excluding zero-volume days, the estimates from the regression in equation (2) would be difficult to interpret. As we explain in our 2003 study, the coefficient of interest from this regression, $\gamma_{i,t}$, captures volume-related return reversals. To capture return reversals that are not volume-related, such as the bid-ask bounce, we include a second independent variable: lagged stock return, $r_{i,d,t}$. However, controlling for a bid-ask bounce is inappropriate on days with zero trading volume because on such days, the stock price reported by CRSP is simply the average of the bid and the ask. The inclusion of an inappropriate control would result in a biased estimate of $\gamma_{i,t}$ because this control, $r_{i,d,t}$, is positively correlated with signed volume $\text{sign}(r^e_{i,d,t}) \cdot v_{i,d,t}$. In other words, if we did not impose $v_{i,d,t} > 0$, the regression in equation (2) would be misspecified. Our imposition of this screen was appropriate, but our failure to mention it in the paper obviously wasn't.

2.4. Modifications of Our Liquidity Measure

Li, Novy-Marx, and Velikov (2017) analyze numerous modifications of our factor construction methodology, such as different approaches to portfolio sorts, different rebalancing frequencies, and different rebalancing months. While these modifications are interesting, our methodological choices seem reasonable a priori. For example, while one could certainly consider other numbers of portfolios, our choice to sort stocks into 10 decile portfolios seems natural. Our choice to allocate an equal number of stocks to each decile portfolio seems equally plausible. Li, Novy-Marx, and Velikov show that allocating stocks differently, by equalizing total market capitalization in each portfolio, produces somewhat weaker results. Yet such "market cap breaks" are not typically used in cross-sectional asset pricing tests, and it is not clear why they should possess more power than the "name breaks" employed in our study. Applying market cap breaks could result in a large imbalance in the number of stocks across portfolios, to the point where some of the decile portfolios could contain a very small number of large stocks.[1] The diversification of idiosyncratic risk within such small portfolios is quite imperfect, making the results noisier.

How often should one rebalance the portfolios of stocks sorted on liquidity betas? We rebalance annually, at the end of the calendar year, which seems perfectly plausible to us. Li, Novy-Marx, and Velikov argue that rebalancing at the monthly frequency would be more natural. We agree that monthly rebalancing would also be plausible, but compared to annual rebalancing, it would significantly increase the portfolio's trading costs. Li, Novy-Marx, and Velikov state in their abstract that their liquidity factor constructed with monthly rebalancing "exhibits significantly weaker performance" than our factor. Yet even under monthly rebalancing, the performance of the liquidity factor remains statistically and economically significant. The full-sample (1968–2015) alphas of the resulting traded factor are all significant, ranging from 0.36% to 0.45% per year, with t-statistics ranging from 2.24 to 2.84 (see

[1] For example, as of this writing, three stocks—Microsoft, Apple, and Amazon—together account for more than 10% of the S&P 500 index. If we were to apply market cap breaks to stocks in this index, we could in principle end up with a decile portfolio containing only two stocks. The CRSP value-weighted market index contains more stocks than the S&P 500, so the problem is somewhat less severe in our context.

Table 3 of Li, Novy-Marx, and Velikov (2017)).

Pontiff and Singla (2019) consider four modifications of our liquidity measure: (1) inclusion of zero-volume days, (2) inclusion of stocks of all price levels, (3) value-weighted index, and (4) zero-intercept restriction. They do not find any of these measures to be priced.

Like our 2003 study, Pontiff and Singla test the joint hypothesis that liquidity risk is priced and that they have the right measure of liquidity. We believe that their rejection of this joint hypothesis is driven by the fact that their four liquidity measures are inadequate.

We plot the four series in Figure 2. The first three series, plotted in Panels A through C, do not appear to be good representations of market-wide liquidity. We base this judgment on a visual examination of these series, invoking the "you know it when you see it" doctrine mentioned in Li, Novy-Marx, and Velikov's introduction. Despite Pontiff and Singla's good intention to improve the estimates of the liquidity series, it seems hard to believe that any of the three series are indeed improvements. On the contrary, they look less appealing as measures of market-wide liquidity. Therefore, the evidence based on these modified series does not imply that liquidity risk is without a premium.

The three series look unappealing not only ex post but also ex ante. The misspecification induced by using zero-volume days is discussed in Section 2.3. The inclusion of stocks with prices below \$5 and above \$1,000 is relatively uncommon. Value-weighting seems less appealing than equal-weighting from the perspective of diversifying away the estimation errors in the $\gamma_{i,t}$ estimates across firms. Additional arguments pertaining to value-weighting are discussed in Section II.C of our 2003 study.

The only series in Figure 2 that exhibits liquidity-like features is the series plotted in Panel D, which corresponds to the zero-intercept modification. Perhaps not surprisingly, the results for that series are more supportive of priced liquidity risk than the results for the other three series. In Panel C of their Table 4, Pontiff and Singla (2019) report that the resulting traded liquidity factor has an average return of 3.13% per year in 1968 through 2017, with a t-statistic of 1.89. While these numbers are lower than the 4.45% return with a t-statistic of 2.67 that Pontiff and Singla report for our baseline index, they still indicate marginal significance of liquidity risk, with a p-value of 0.06.

More important, we are not convinced by the argument that imposing the zero-intercept restriction improves our procedure. It is true that imposing this restriction saves a degree of freedom, but it also introduces misspecification. Pontiff and Singla argue that the restriction is reasonable because it generates one-day-ahead forecasts of individual stock returns that are equal to market returns when the independent variables in regression (2) are zero. While market returns are indeed plausible unconditional forecasts of individual stock returns, it is not clear that they are plausible conditional on the very unlikely event that the independent variables—lagged return and signed volume—are both equal to zero. Zeroing out the intercept could result in biased regression estimates. We think of the zero-intercept-modification series as a slightly misspecified version of our baseline series. Given that, we do not find it surprising that this series produces slightly weaker results.

Besides the four modifications plotted in Figure 2, Pontiff and Singla consider five additional alternative liquidity measures: (1) the proportion of zero returns, (2) the Amihud index, (3) estimated relative bid-ask spread, (4) a hybrid measure, and (5) a double variation on our original index. They find that none of these measures are priced. Again, the joint hypothesis problem mentioned earlier rears its head. Pontiff and Singla's findings do not imply that liquidity risk is not priced; they could simply mean that the five series are poor time-series proxies for aggregate liquidity. All five series are plotted in Figure 3. Their visual examination indeed casts doubt on whether any of them capture liquidity. Most of these series fail to pick up well-known liquidity-crisis months. For example, with the exception of the bid-ask spread, all other alternative series fail to pick up the undisputed evaporation of liquidity during the 2008 financial crisis. At the same time, the alternative series do pick up months that are not known for being low-liquidity based on other sources. These series exhibit both Type I and Type II errors, so to speak.

The only series whose plot reveals some resemblance to liquidity is the bid-ask spread. While this spread could be a good measure of liquidity from the perspective of small retail investors, who might trade at the bid and ask quotes, it is less useful to large investors, whose trades have price impact. We measure liquidity by the temporary price impact resulting from a given amount of order flow. The bid-ask spread cannot capture price impact because it does not take the market's depth into account. The correlation between the levels of our series and the bid-ask spread in August 1962 through December 2017 is 29.6%. The two series clearly capture different dimensions of liquidity. Not all dimensions of liquidity need to be priced. If asset prices are determined primarily by large investors, our liquidity measure, which aims to capture price impact, is more likely to be priced than bid-ask spread.

2.5. Momentum

Pástor and Stambaugh (2003) find that the liquidity risk factor accounts for half of the profits of the momentum strategy in stocks. Subsequent studies provide further evidence of a link between liquidity risk and momentum. Sadka (2006) also finds that a substantial portion of the profit in stock momentum is explained by liquidity risk. To estimate the latter, he constructs an alternative measure of market-wide liquidity using intraday data. Asness, Moskowitz, and Pedersen (2013) conclude that a link between momentum and liquidity risk exists quite broadly. Using multi-country data, they find that, within stocks as well as other asset classes, momentum returns exhibit positive sensitivities to measures of market-wide liquidity.

Li, Novy-Marx, and Velikov (2017) argue that our finding of a link between momentum and liquidity risk is due to the inclusion of momentum-related variables, namely past return and share price, among the predictors of liquidity betas. They note that if the predicted-beta liquidity factor is constructed by excluding cumulative return and stock price, it can no longer explain much of momentum profit. That is useful to know. However, we did not include these two variables in order to explain momentum. We included them because they show up as significant determinants of liquidity betas (see Table 2 in Pástor and Stambaugh (2003)). Excluding a significant determinant would cause an omitted-variable problem in the analysis. For that reason, we believe it makes sense to include both cumulative return and stock price in the construction of the predicted-beta liquidity factor.

Li, Novy-Marx, and Velikov (2017) also argue that explaining half of momentum is one of our most important results. We do not share that view. Our main results instead relate to the measurement of market-wide liquidity and the pricing of liquidity risk more generally. Our objective was not to investigate momentum, which is not central to our study. Evidence of a momentum link emerged simply when investigating the performance of traded liquidity factors in the context of other popular traded factors.

3. How to Use the Liquidity Factors

3.1. Traded and Non-Traded Factors

Our liquidity factor comes in two versions: traded and non-traded. The non-traded factor, $\mathcal{L}_t$, is in equation (8) of Pástor and Stambaugh (2003). The traded factor, LIQ, is the payoff on the 10–1 portfolio that is long stocks with the highest historical liquidity betas and short stocks with the lowest historical liquidity betas. We are sometimes asked by colleagues which of the two factors they should use. The answer depends on the objective of the analysis.

For most objectives, the appropriate liquidity factor is the non-traded one, $\mathcal{L}_t$. This is the primary liquidity factor, designed to capture innovations in market liquidity. If one wishes to estimate an asset's liquidity risk, one should use $\mathcal{L}_t$.

The traded factor, LIQ, is useful in estimating alpha with respect to a multifactor model that includes a role for liquidity risk. In order to interpret the intercept from a factor model regression as alpha, all factors must be payoffs on tradable positions. Because $\mathcal{L}_t$ is not traded, it would be inappropriate to interpret the intercept β_i^0 from the regression in equation (1) as alpha. However, if we replace $\mathcal{L}_t$ in equation (1) by LIQ, the intercept can be interpreted as an alpha, because all four factors are then traded.

Compared to other traded factors in popular multifactor models, LIQ is rather unique in having an identified non-traded factor underlying it. Moreover, our traded factor is formed by sorting on risk estimates, unlike other long-short factors such as SMB and HML of Fama and French (1993, 2015). The latter factors are often described as returns on mimicking portfolios, but the non-traded factors (state variables) that those portfolios presumably mimic are not specified. That slack cannot be afforded to LIQ.

One definition of a mimicking portfolio is advanced by Huberman, Kandel, and Stambaugh (1987), hereafter HKS. In their setting, portfolios "mimic" non-traded factors if betas with respect to the former can be used instead of betas with respect the latter to explain the cross-section of expected returns. HKS provide the general characterization for the weights in mimicking portfolios. One admissible specification constructs portfolios maximally correlated with the non-traded factors. As HKS explain, however, many other constructions are also admissible mimicking portfolios. Our traded factor, LIQ, is not designed to maximize its correlation with innovations in liquidity. Indeed, the simple correlation between LIQ and $\mathcal{L}_t$ is only 5.3%, based on Pontiff and Singla's data in the longest possible time period January 1932 through December 2017.

Using LIQ in a model with traded factors has a simple investment-based motivation, as we explain in Section IV of Pástor and Stambaugh (2003). Adding LIQ to an investment opportunity set consisting of other popular factors substantially increases the maximum obtainable Sharpe ratio. Therefore, other assets are less likely to produce alphas with respect to factors that include LIQ. If LIQ reduces alphas, then betas with respect to LIQ help explain expected returns. Thus LIQ behaves as a mimicking portfolio in the sense of HKS, though it is not constructed to obey the formal conditions derived by that study, which does not address empirical implementation. Instead, LIQ has a simple long-short construction typical of other traded factors. Therefore, except when a traded factor is needed, one should use the underlying non-traded factor, $\mathcal{L}_t$, rather than its imperfect stand-in, LIQ.

3.2. Aggregate versus Firm-Level Liquidity

To construct our measure of aggregate market liquidity in month t, we average the slope estimates $\hat{\gamma}_{i,t}$ across stocks i (see equations (1) and (5) of Pástor and Stambaugh (2003)). One might be tempted to use the individual-firm estimates $\hat{\gamma}_{i,t}$ on their own as firm-level measures of liquidity. However, that is not an appropriate use of our liquidity measure. As we explain in Section III.D of Pástor and Stambaugh (2003), the $\hat{\gamma}_{i,t}$ estimates are too noisy to be useful at the individual firm level. Despite our warnings, several studies, such as Goyenko, Holden, and Trzcinka (2009), examine the cross-section of the $\hat{\gamma}_{i,t}$ estimates. Not surprisingly, they find that these estimates are noisy. We do not use the firm-level estimates $\hat{\gamma}_{i,t}$ for any purpose other than to construct a market-wide measure of liquidity. When we average these estimates, much of the noise diversifies away, resulting in what we believe is an appealing measure of market-wide liquidity.

3.3. Historical versus Predicted Liquidity Betas

In our 2003 study, we produce two versions of the traded liquidity factor. Both versions go long high-liquidity-beta stocks and short low-liquidity beta stocks, but they differ in how the betas are estimated. One version uses "historical" liquidity betas, which are estimated values of $\beta_i^{\mathcal{L}}$ from regression (1) estimated over the previous 60 months of data. The other version uses "predicted" liquidity betas, which are estimated as a linear function of seven firm-level characteristics (see equation (10) in Pástor and Stambaugh, 2003). The results based on both versions of the traded factor suggest that liquidity risk is priced.

In our 2003 study, we put a larger emphasis on predicted betas because the associated procedure produced a larger spread in post-ranking liquidity betas. Specifically, we sort stocks into portfolios in two different ways: based on pre-ranking predicted betas and based on pre-ranking historical betas. At the time when we wrote our 2003 study, the sort on predicted betas produced a larger spread in post-ranking liquidity betas. Comparing Tables 3 and 7 in Pástor and Stambaugh (2003), the 10-1 portfolio's post-ranking liquidity beta is 8.23 ($t = 2.37$) when the portfolios are formed based on predicted betas, but it is only 5.99 ($t = 1.88$) when the portfolios are formed based on historical betas.

Years later, we noticed that the beta pattern flipped in the expanded sample: historical betas now do a better job than predicted betas in creating dispersion in the post-ranking betas. For example, when we extend Tables 3 and 7 through December 2018, the 10-1 portfolio's post-ranking liquidity beta is only 3.24 ($t = 0.91$) when the portfolios are formed based on predicted betas, but it is 8.04 ($t = 3.16$) when the portfolios are formed based on historical betas. It appears that our model for predicting liquidity betas is somewhat unstable over time. To maintain the quality of our traded liquidity factor, we decided about ten years ago to focus on historical betas, and we have been doing that ever since. A side benefit of this re-focus is that historical betas are substantially easier to estimate than predicted betas. Estimating historical betas requires only CRSP data, whereas estimating predicted betas requires also Compustat data.

Li, Novy-Marx, and Velikov (2017) report some difficulty in replicating our traded factor based on predicted betas. We are not sure what the source of this difficulty is. Given our present focus on historical betas, we have decided not to allocate time to this issue. It is comforting that despite its nontrivial differences from our factor, the predicted-beta-based liquidity factor constructed by Li, Novy-Marx, and Velikov earns positive alphas both in and out of sample. Their Table 5 shows that the factor performs most strongly against the five-factor Fama-French model, with alphas that are positive and significant both in and out of sample (0.59% per month ($t = 4.04$) in 1966–1999 and 0.73% per month ($t = 2.56$) in 2000–2015). That is, even though their predicted-beta-based traded liquidity factor differs somewhat from ours, both factors appear to be priced.

4. Asymmetric Liquidity Shocks: Implications

The liquidity factor sleeps most of the time but wakes up during financial crises and other significant disruptive events. This asymmetric nature of liquidity shocks, which is apparent from Figure 1, has two important implications.

First, liquidity risk can explain, at least in principle, why asset return correlations tend to rise sharply during crises. Most risky assets have positive exposures to liquidity fluctuations: their prices tend to fall when liquidity dries up. In normal times, these positive liquidity exposures contribute little to asset returns because they are multiplied by small liquidity shocks. However, during crises, large negative realizations of the liquidity factor are multiplied by positive liquidity betas to produce large negative contributions to asset returns. As a result of this common exposure to liquidity risk, the prices of many risky assets fall and their correlation rises sharply. Formally, in equation (1), $\beta_i^{\mathcal{L}} \mathcal{L}_t \approx 0$ when $\mathcal{L}_t \approx 0$, but $\beta_i^{\mathcal{L}} \mathcal{L}_t$ is large and negative for most assets when $\mathcal{L}_t << 0$, resulting in highly correlated asset returns during liquidity crises.

Second, liquidity betas are relatively hard to estimate, especially in normal times. Running a regression of asset returns on liquidity shocks, whose values in normal times are close to zero, produces noisy estimates of liquidity betas. To achieve some precision, it helps to estimate liquidity betas over periods that include substantial shocks to liquidity, such as financial crises. For example, liquidity betas estimated over a period that includes the 2008 financial crisis are likely to be more precise than those estimated in normal times. That the 2008 crisis occured after our initial sample period could be the reason why the post-sample evidence on liquidity risk is even stronger than the in-sample evidence. Another way of improving the precision of liquidity beta estimates is to form portfolios. To illustrate these implications of skewed liquidity shocks, we conduct a simulation analysis.

4.1. Simulation Analysis

We assume that excess stock returns are determined by the two-factor model

$$r_{i,t} = \beta_i^{LIQ} F_t^{LIQ} + \beta_i^{MKT} F_t^{MKT} + \epsilon_{i,t} \ , \tag{3}$$

where F_t^{LIQ} and F_t^{MKT} are the realizations of two hypothetical factors that we refer to as liquidity and market factors, respectively. The liquidity factor's realizations exhibit negative skewness, whereas the market factor's realizations are symmetrically distributed.

We simulate monthly time series of both factors, which we make uncorrelated with each other. Each factor realization is the sum of a constant premium and a zero-mean innovation. We assign a premium of 0.33% per month, or 4% per year, to each factor. We assume both factors contribute equally to the volatility of stock returns and that the volatility of $F_t^{LIQ} + F_t^{MKT}$ is 20% per year, implying a monthly volatility of 4.08% for each factor. We draw each monthly realization of F_t^{MKT} from a normal distribution with a mean of 0.33% and standard deviation of 4.08%. We construct monthly realizations of the liquidity factor as $F_t^{LIQ} = S_t + J_t$, where S_t represents "small" shocks and J_t captures occasional jumps. Specifically, S_t is normally distributed with a mean of 0.33% and a standard deviation of 1.22%, whereas $J_t = -30\%$ when a jump occurs and $J_t = 0$ otherwise. We set the probability of a jump in any given month to 1/60.

Figure 4 plots simulated innovations of both factors over a 50-year period. Panel A shows that the liquidity shocks are strongly negatively skewed, with eight jumps ("liquidity crises") occurring in this typical simulated history. These asymmetric dynamics of simulated liquidity shocks are reminiscent of actual liquidity realizations, plotted in Figure 1. In contrast, Panel B shows a symmetric pattern of simulated market shocks resembling that of actual stock market returns.

We assume that stocks' liquidity betas and market betas, β_i^{LIQ} and β_i^{MKT}, are both uniformly and independently distributed within the range of 0.5 to 1.5. For each individual stock, we assume that the two factors together explain 30% of the variance of the stock's return, $r_{i,t}$. Given the stock's betas, this assumption implies the value for the variance of the stock's $\epsilon_{i,t}$, which is normally distributed, independently across stocks and months. The draws of $\epsilon_{i,t}$ are combined with the stock's betas and the draws of F_t^{LIQ} and F_t^{MKT} to construct $r_{i,t}$ as in equation (3). Fixing the sample length at 60 months, we estimate equation (3) as an OLS regression with an intercept. For each simulated sample, we collect the sample estimates of β_i^{LIQ} and β_i^{MKT}, denote them by $\hat{\beta}_i^{LIQ}$ and $\hat{\beta}_i^{MKT}$, and construct the estimation errors

$$e_i^{LIQ} = \hat{\beta}_i^{LIQ} - \beta_i^{LIQ} \tag{4}$$

$$e_i^{MKT} = \hat{\beta}_i^{MKT} - \beta_i^{MKT} . \tag{5}$$

We simulate many 60-month samples of factors and returns on individual stocks. Panel A of Figure 5 plots the distributions of e_i^{LIQ} and e_i^{MKT} across the simulated samples. There is a stark difference between the two distributions, especially in the tails. The distribution of e_i^{MKT} tails off around the values of -1 and 1, whereas that of e_i^{LIQ} tails off around -3 and 3. In other words, it is extremely rare for the estimates of market betas to be off by more than 1, whereas the estimation errors in liquidity beta can be as large as 3. In that sense, liquidity betas are clearly harder to estimate than market betas.

The difficulty in estimating liquidity betas stems from the rarity of liquidity crises. The betas are especially hard to estimate in "calm" periods in which no crises occur. To establish this result, we divide our simulated samples into two subsets: those containing at least one large liquidity shock (i.e., at least one non-zero value of J_t) and those with no large shocks. The first subset includes about 64% of all simulated samples. Panel A of Figure 6 shows that the distribution of e_i^{LIQ} is substantially wider in samples without large liquidity shocks. In such samples, we sometimes observe estimation errors as large as 3 in magnitude, whereas errors larger than 1 are extremely rare in samples with at least one large liquidity shock. Therefore, a simple way to improve the precision of liquidity beta estimates is to compute them over periods that include at least one liquidity crisis.

Another effective way to reduce the estimation error in liquidity betas is to form portfolios. To illustrate this point, we expand our simulation exercise by adding portfolio formation. Following the same procedure as before, we simulate many 60-month samples of factor and return realizations for 2,000 individual stocks. For each sample, we sort stocks into 10 portfolios by their true liquidity betas, and also separately by their true market betas. We compute the liquidity (market) beta of each portfolio as the average of the liquidity (market) betas across all stocks in that portfolio, for both true and estimated betas.

Panels B of Figures 5 and 6 plot the distributions of the estimation errors in portfolio betas. Compared to Panels A of the same figures, the distributions are narrower by an order of magnitude, indicating that beta estimates are much more precise when computed at the portfolio level rather than the individual stock level. At both levels, the estimation errors in liquidity betas are larger than those in market betas because the underlying problem—the scarcity of liquidity crises—cannot be completely eliminated by forming portfolios. Nonetheless, these plots show that portfolio formation is especially valuable when estimating liquidity betas, given the large amount of noise in their individual-stock estimates.

5. Conclusion

Twenty years after LTCM, our profession has a deeper understanding of liquidity risk and its effects on asset prices. Hundreds of studies have examined the pricing of liquidity risk in various markets and the exposures of professional money managers to this risk. This literature has also been among the few beneficiaries of the 2008 financial crisis.

We commend Li, Novy-Marx, and Velikov (2017) and Pontiff and Singla (2019) for their thorough and successful replication of our 2003 study. In addition to handling their task professionally and gracefully, they offer numerous interesting insights on this topic. They find the liquidity risk factor to be priced not only in sample but also post sample, and the post-sample performance they find is even stronger than the in-sample performance. They also report that various alternative liquidity measures are not priced, which leads them to voice skepticism about the cross-sectional relation between liquidity risk and expected stock returns. However, many of these alternative measures do not resemble liquidity when plotted, while others may not be capturing the price-relevant dimension of liquidity. Upon reviewing their evidence, we continue to believe it is reasonable to infer that liquidity risk is priced.

Besides addressing replication, we also offer recommendations regarding the usage of our liquidity factor, such as when to use its traded versus non-traded version and whether to use historical or predicted liquidity betas. We explain why liquidity betas are inherently difficult to estimate and offer tips on how to improve their precision.

Despite the immense growth in research on liquidity risk, important questions remain unanswered. For example, what predicts the periodic seizures in market liquidity? Our work shows how one can measure market-wide liquidity and provides clues about how to hedge it, but not how to predict it. The origins of liquidity shocks are of obvious interest not only to investors but also to policymakers.

The replication initiative of the *Critical Finance Review* raises a number of interesting questions about replication in general. For example, should the replications that are published be a subset of those conducted, and, if so, how should the former be selected? Galiani, Gertler, and Romero (2017) report evidence that editors are more likely to publish replication studies that overturn previous results rather than support them. Is such a preference desirable? If journals designate ex ante the studies to be replicated, should editors select studies that they suspect are more likely to be overturned, or should they instead select randomly from studies deemed influential by an objective standard? The latter selection could be more informative, for example, about the profession's error rate. If a study goes beyond replicating the original study and introduces additional results from alternative specifications or sample periods, who replicates the replicators? If replicators try enough variations, they will almost surely find some by chance that fail to reproduce the original result. What are the incentives of a replicator? It is increasingly common to make the computer code underlying the paper's analysis publicly available. Could such public posting jeopardize the independence of code written for subsequent replications? These questions, and many others, are open for debate in our academic community.[2]

[2] All of these questions apply to replications in general. None of them should reflect in any way on the two replications of our study, which we believe were conducted with the utmost care and integrity.

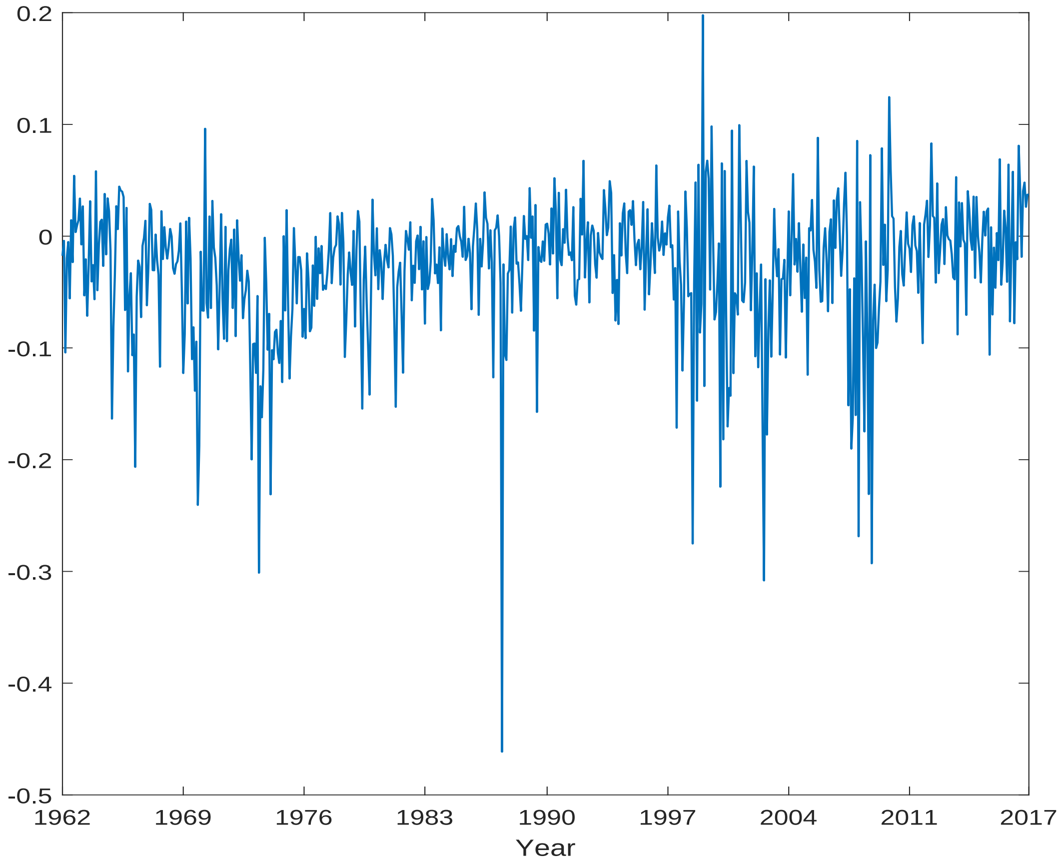


Figure 1. The Original PS Liquidity Series. This figure plots the aggregate liquidity series of Pástor and Stambaugh (2003) extended through 2017 by Pontiff and Singla (2019).

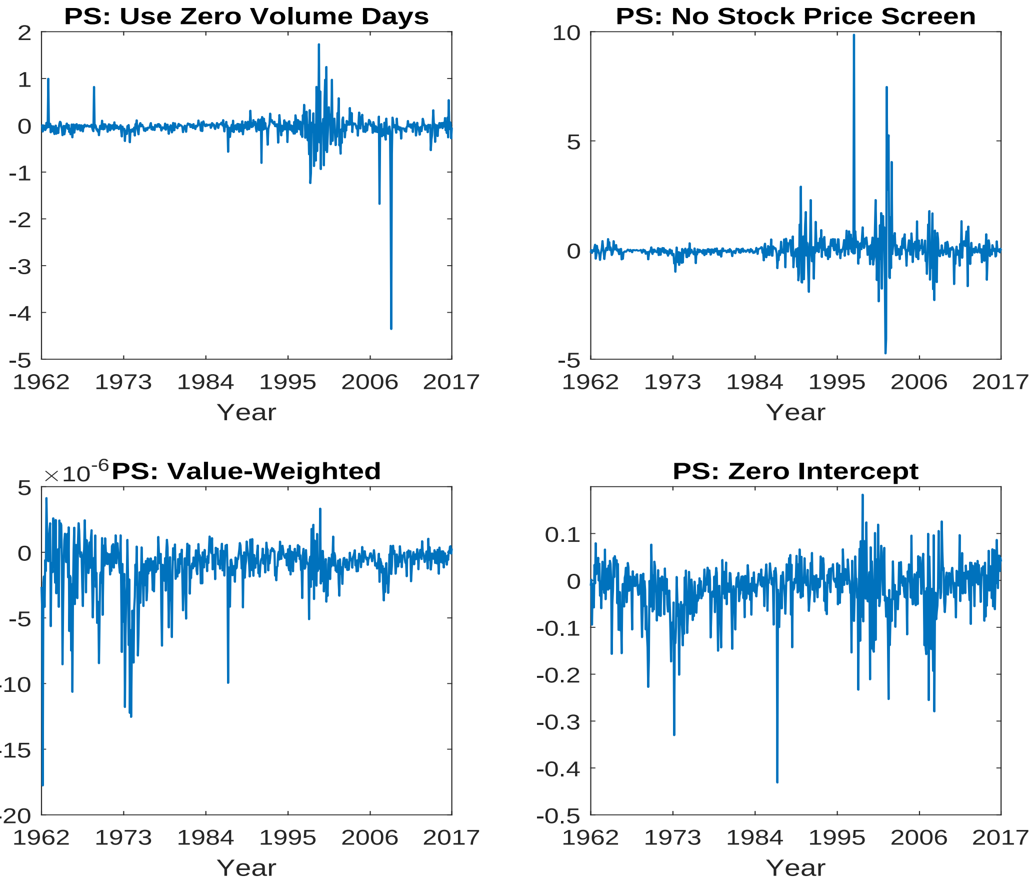


Figure 2. Modifications of the PS Liquidity Series. This figure plots the four modifications of the Pástor and Stambaugh (2003) aggregate liquidity series constructed by Pontiff and Singla (2019).

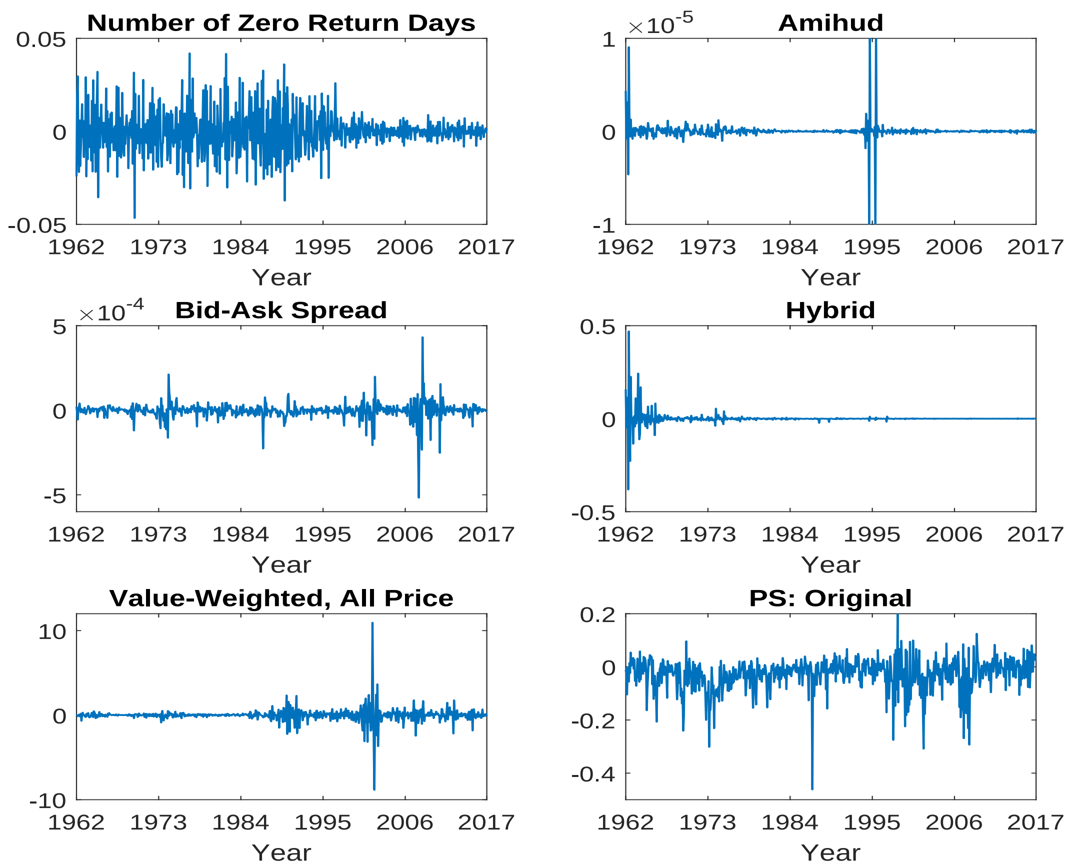
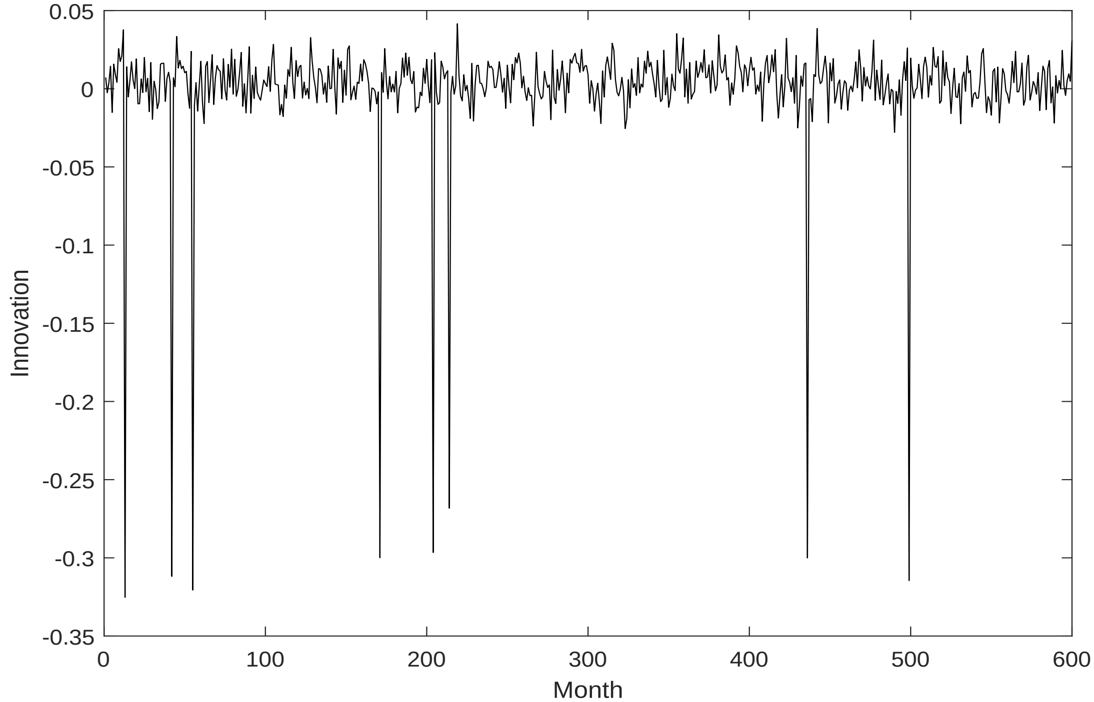


Figure 3. Alternative Liquidity Series. Panel A plots the original Pástor-Stambaugh aggregate liquidity series. The remaining panels plot the five alternative liquidity series constructed by Pontiff and Singla (2019).

Panel A. Simulated innovations: Liquidity factor



Panel B. Simulated innovations: Market factor

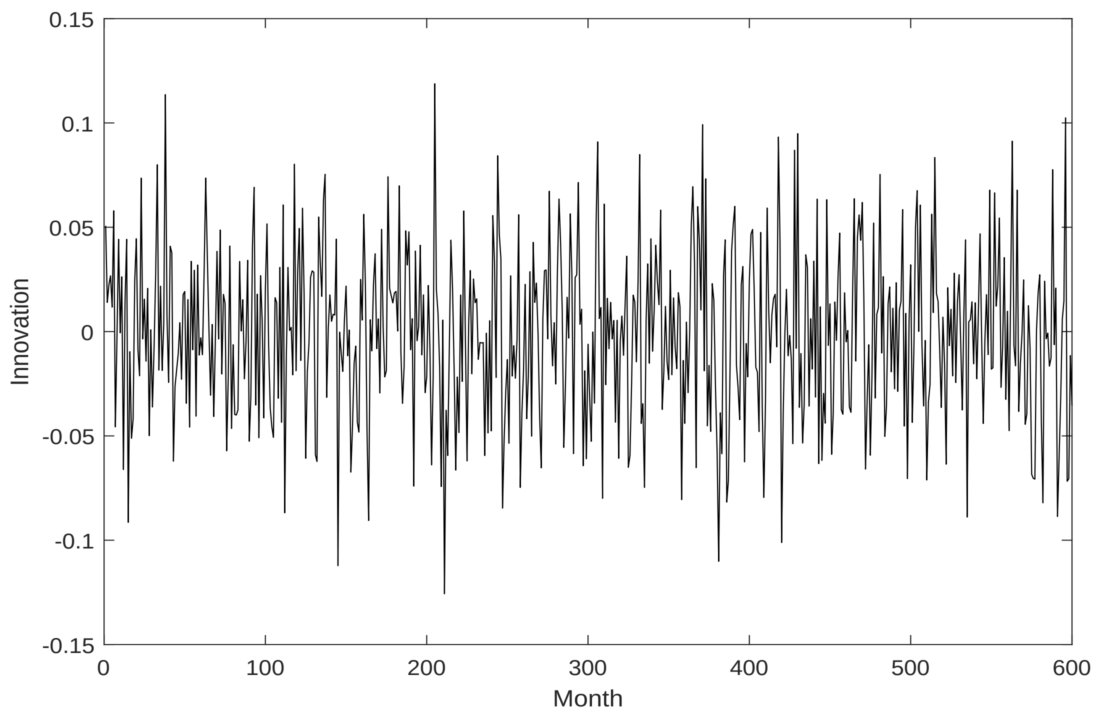
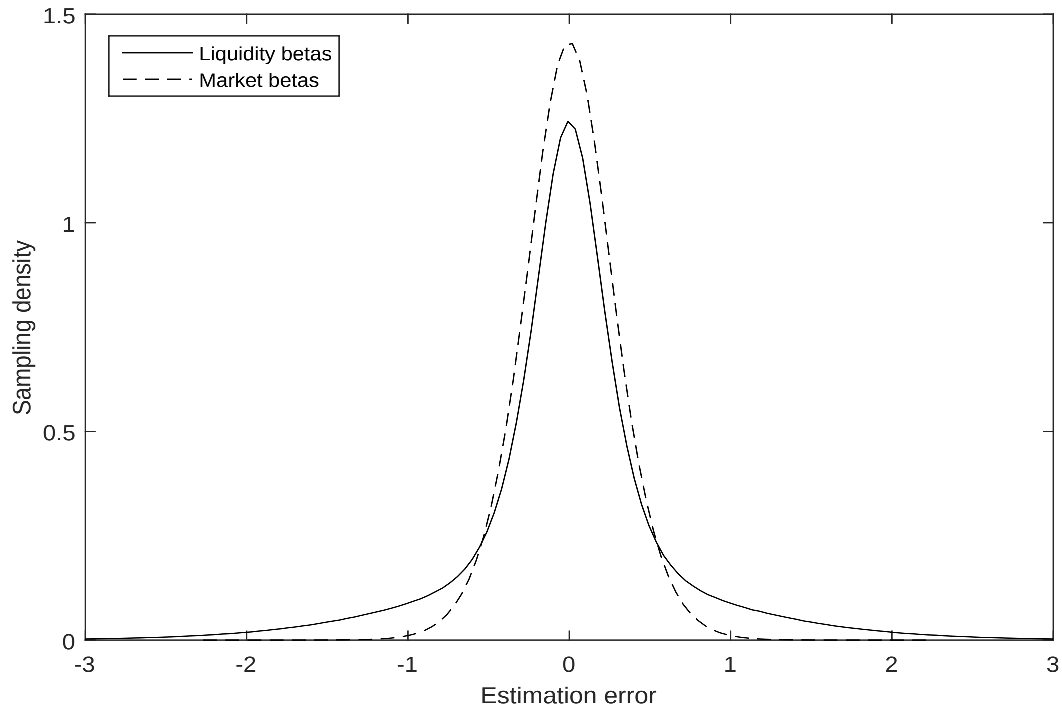


Figure 4. Simulated innovations for a hypothetical sample. For one simulated 50-year sample, Panel A displays the liquidity innovations and Panel B displays the market innovations.

Panel A. Individual stocks



Panel B. Portfolios

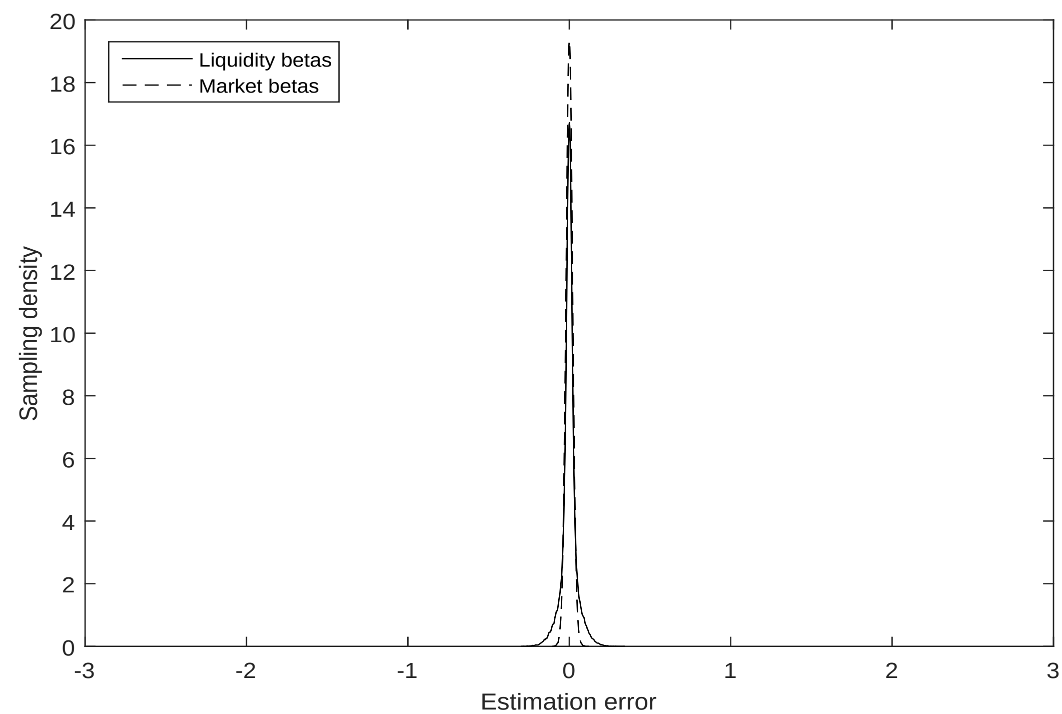


Figure 5. Distributions of beta estimation errors. The plots display simulation-based sampling densities of estimation errors in liquidity betas and market betas estimated with five years of monthly data. Panel A displays the densities of errors in betas estimated for individual stocks; Panel B displays the densities of errors in estimated portfolio betas.

Panel A. Individual stocks

Panel B. Portfolios

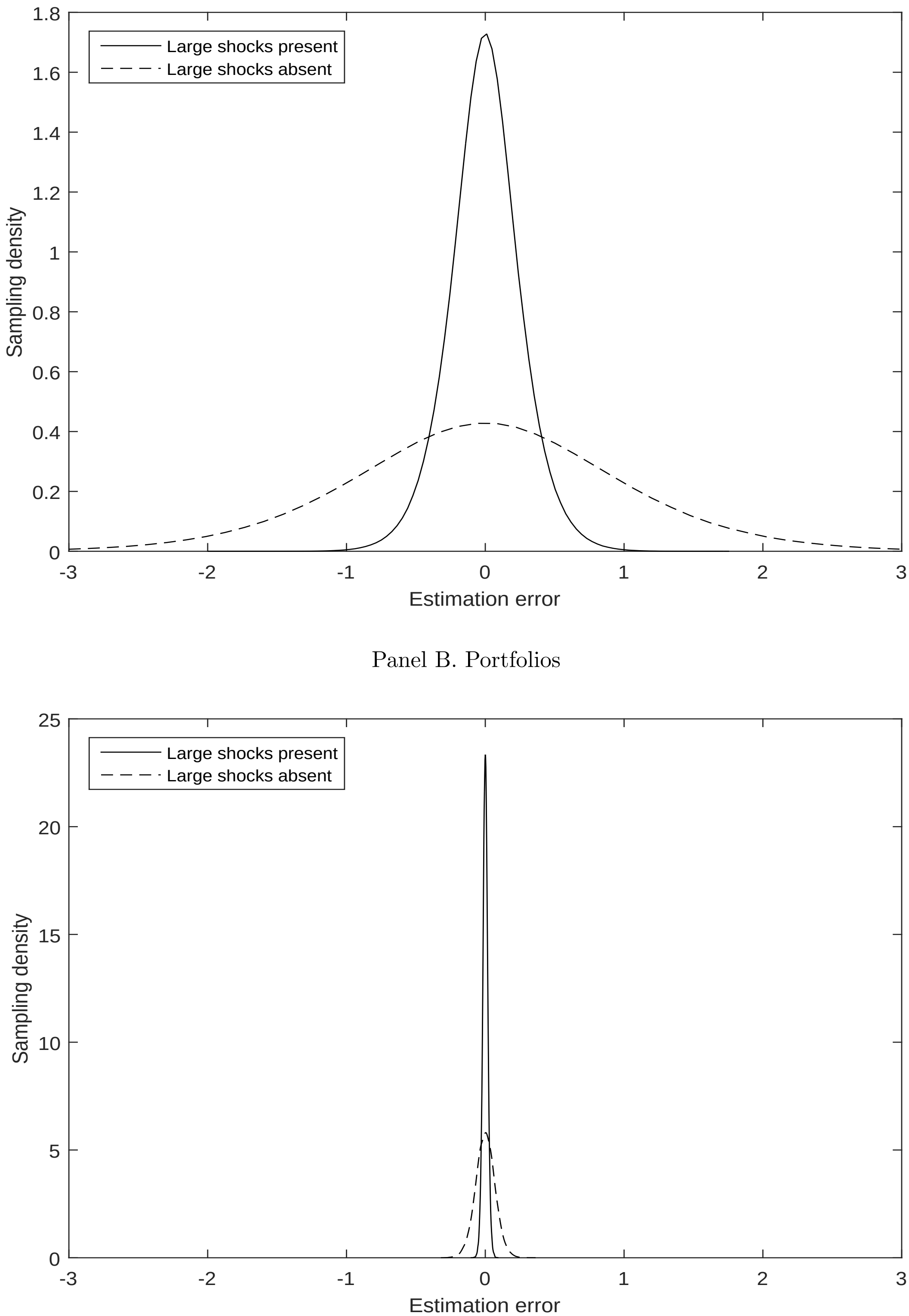


Figure 6. Effect of liquidity shocks on liquidity beta estimates. The plots display simulation-based sampling densities of errors in liquidity betas estimated during five-year periods with and without large liquidity shocks. Panel A displays the densities of errors in betas estimated for individual stocks; Panel B displays the densities of errors in estimated portfolio betas.

REFERENCES

Acharya, Viral V., and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad, 2007, Liquidity and expected returns: Lessons from emerging markets, *Review of Financial Studies* 20, 1783–1831.

Bongaerts, Dion, Frank De Jong, and Joost Driessen, 2017, An asset pricing approach to liquidity effects in corporate bond markets, *Review of Financial Studies* 30, 1229–1269.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

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

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Franzoni, Francesco, Eric Nowak, and Ludovic Phalippou, 2012, Private equity performance and liquidity risk, *Journal of Finance* 67, 2341–2373.

Galiani, Sebastian, Paul Gertler, and Mauricio Romero, 2017, Incentives for replication in economics, NBER Working Paper No. 23576, Cambridge, MA.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka, 2009, Do liquidity measures measure liquidity? *Journal of Financial Economics* 92, 153–181.

Huberman, Gur, Shmuel Kandel, and Robert F. Stambaugh, 1987, Mimicking portfolios and exact arbitrage pricing, *Journal of Finance* 42, 1–9.

Li, Hongtao, Robert Novy-Marx, and Mihail Velikov, 2017, Liquidity risk and asset pricing, *Critical Finance Review*, forthcoming.

Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Pontiff, Jeffrey, and Rohit Singla, 2019, Liquidity risk?, *Critical Finance Review*, forthcoming.

Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80, 309–349.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge-fund returns, *Journal of Financial Economics* 98, 54–71.