June 21, 2016

The Dynamics of Market Efficiency

Dominik M. Rösch, Avanidhar Subrahmanyam, and Mathijs A. van Dijk

Rösch is at the State University of New York, Buffalo; Subrahmanyam is at the UCLA Anderson School; van Dijk is at the Rotterdam School of Management, Erasmus University. E-mail addresses: drosch@buffalo.edu, subra@anderson.ucla.edu, and madijk@rsm.nl, respectively. We thank Stefan Nagel (the editor), two anonymous referees, Yakov Amihud, Tarun Chordia, Carole Comerton-Forde, Thierry Foucault, Amit Goyal, Terry Hendershott, Craig Holden, Sreeni Kamma, Andrew Karolyi, Ed Lin, Marc Lipson, Steve Mann, Christophe Pérignon, Veronika Pool, Vikas Raman, Raghu Rau, Matti Suominen, Kumar Venkataraman, Avi Wohl, Hong Yan, and participants at the 2012 Brazilian Finance Conference (São Paulo), the 2012 EFMA meetings (Barcelona), the 2012 Frontiers of Finance Conference (Warwick Business School), the 2013 Campus for Finance conference (WHU Otto Beisheim School of Management), the 2013 EFA meetings (Cambridge), and at seminars at Deakin University, Erasmus University, Goethe University Frankfurt, Indiana University, UCLA Anderson, University of Cambridge, University of Manchester, and University of South Carolina for valuable comments. This work was carried out on the National e-infrastructure with the support of SURF Foundation. We thank SURFsara, and in particular Lykle Voort, for technical support on computing and storage, and OneMarketData for the use of their OneTick software. Van Dijk gratefully acknowledges financial support from the Vereniging Trustfonds Erasmus Universiteit Rotterdam and from the Netherlands Organisation for Scientific Research through a "Vidi" grant.

Abstract

The Dynamics of Market Efficiency

This paper studies the dynamics of high-frequency market efficiency measures. We provide evidence that these measures co-move across stocks and with each other, suggesting the existence of a systematic market efficiency component. In vector autoregressions, we show that shocks to funding liquidity (the TED spread), hedge fund assets under management, and a proxy for algorithmic trading are significantly associated with systematic market efficiency. Thus, stock market efficiency is prone to systematic fluctuations, and, consistent with recent theories, events and policies that impact funding liquidity can affect the aggregate degree of price efficiency.

In a financial market that is relatively free of frictions and of high quality (i.e., one that is "efficient"), prices accurately reflect fundamentals, and, in doing so, obey the law of one price that assets with identical cash flows sell for the same price. For most of its life, the finance profession has treated market efficiency as a static concept. The seminal taxonomy in Fama (1970) of weak-, semi-, and strong-form efficiency inspires debate on which of these best describes financial markets, but it does not allow for market efficiency itself to vary through time, in predictable as well as unexpected ways. And yet, of course, there are sound reasons to expect such dynamic behavior. Market efficiency is governed by arbitrage activity and market making capacity, both of which facilitate convergence of prices to their efficient market benchmarks. In turn, the efficacy of arbitrage and market making is influenced by financial frictions (such as limited capital, transaction costs, short-sale constraints, and idiosyncratic volatility)[1] whose severity varies considerably over time.

The finance literature has developed a number of distinct measures to capture the degree of efficiency. For example, traditional measures that test whether stock prices follow a random walk (e.g., variance ratios, intraday return predictability) date back to Fama (1970). Alternative measures are based on predictable intraday patterns in the cross-section of stock returns (e.g., Heston, Korajczyk, and Sadka, 2010) or directly measure the pricing error relative to the efficient price (e.g., Hasbrouck, 1993). Yet other measures assess the extent to which markets obey the law of one price (such as put-call parity deviations; e.g., Finucane, 1991; Cremers and Weinbaum, 2010).

The above measures of stock price efficiency have largely been investigated separately in the literature. However, we note that they all are intimately linked to arbitrage and market making, which are impeded by time-varying financial frictions. And although

[1]See, for example, Shleifer and Vishny (1997), Mitchell, Pulvino, and Stafford (2002), and D'Avolio (2002) for theoretical and empirical explorations of how limits to arbitrage can cause market inefficiencies to persist.

these frictions differ across individual securities[2], they also have a systematic component.[3] Thus, there may be a significant systematic component to the time-varying behavior of market efficiency measures.

Motivated by the above observations, in this paper, we ask the following questions. To what extent do different market efficiency measures vary over time? Do different market efficiency measures co-move across stocks as well as with each other? And, if there is evidence of a systematic market efficiency component across stocks and across measures, what are the economic forces (such as funding liquidity or other factors that affect the efficacy of arbitrage) that drive it? These questions are relevant since investors, exchange officials, and policy-makers should care about whether the efficiency of financial markets is prone to fluctuation in a systematic way, and about what factors influence such systematic variation. For example, investors' allocations to equities may be influenced by their systematic degree of price efficiency. Moreover, researchers could benefit from a better understanding of the extent to which the different efficiency measures used in the literature are related, and of whether they can be used as substitutes.

To address these questions, we first compute daily market efficiency estimates for individual stocks based on four efficiency measures that are widely used and that can be computed by stock-day for a large sample of stocks: intraday return predictability based on past order flow or past returns (Boehmer and Wu, 2007; Andrade, Chang, and Seasholes, 2008), variance ratios (Lo and MacKinlay, 1989; Bessembinder, 2003),

[2]See, e.g., Benston and Hagerman (1974) and Nagel (2005) for evidence on cross-sectional variation in stock-level illiquidity and short-sales constraints, respectively.

[3]See, e.g., Hasbrouck and Seppi (2000) and Chordia, Roll, and Subrahmanyam (2000) for evidence on systematic variation in market liquidity across stocks. In addition, time-variation in liquidity depends on variables that influence market making behavior, such as market volatility and net order imbalances (Chordia, Roll, and Subrahmanyam, 2002) as well as macroeconomic funding constraints (Brunnermeier and Pedersen, 2009). Gârleanu and Pedersen (2011) link deviations from the law of one price to variation in the aggregate shadow cost of capital. Desai, Ramesh, Thiagarajan, and Balachandran (2002), Jones and Lamont (2002), and Asquith, Pathak, and Ritter (2005) provide evidence of systematic variation in short-sale constraints.

intraday Hasbrouck's (1993) pricing errors, and put-call parity deviations in the corresponding options markets (Finucane, 1991; Cremers and Weinbaum, 2010). We compute these measures using all NYSE stocks over an extended sample period of fifteen years (using 14.3 billion transactions). We show that all measures exhibit substantial time-variation.[4] We construct market-wide measures of efficiency from each of the stock-level measures and estimate the degree of "co-movement in efficiency" as the R^2 from regressions of stock-level measures on market-wide measures. These analyses show that time-variation in efficiency measures has a material common component across stocks, which indicates that the market efficiency measures are prone to systematic improvement and deterioration.

We then examine co-movement in aggregate market efficiency across measures by estimating correlations across the different monthly, market-wide efficiency measures. In this analysis, we also include the market-wide cross-sectional return predictability measure proposed by Heston, Korajczyk, and Sadka (2010). These correlations are mostly economically substantial and statistically significant, with the notable exception of the correlations of the variance ratio measure vis-à-vis the other efficiency measures. This finding indicates that four of the five market-wide efficiency measures share significant common variation, which suggests the existence of a systematic market efficiency component across stocks and measures. We extract the component via principal component analysis from the monthly time-series of these four market-wide efficiency measures and show that this first component explains almost 40% of their joint variation.

Our next goal is to analyze the economic forces that drive the dynamics of systematic market efficiency. An expanding body of theoretical research emphasizes the importance of funding constraints as a friction that hampers arbitrage (e.g., Shleifer and Vishny, 1997;

[4]We ensure that our results on (systematic) variation in efficiency are not driven by underlying (systematic) variation in stock-level (il)liquidity by orthogonalizing our stock-level efficiency measures with respect to stock-level liquidity before running any further analyses.

Gromb and Vayanos, 2002; Brunnermeier and Pedersen, 2009; Gârleanu and Pedersen, 2011; Liu and Mello, 2011). Building on these studies, we hypothesize that variation in funding liquidity and the overall intensity of arbitrage activity affect the different efficiency measures for many stocks at the same time, and thus the systematic market efficiency component.

To study the determinants of time-variation in the systematic market efficiency component, we use it as the main variable of interest in vector autoregressions (VARs). As other endogenous variables, we include the TED spread (a common indicator of funding liquidity), hedge fund assets under management (a proxy for the amount of capital available for arbitrage activity), and the total number of quote updates divided by aggregate dollar trading volume (a proxy for algorithmic trading, inspired by Boehmer, Fong, and Wu, 2015). We also include market volatility as another potentially important determinant of the efficacy of market making and arbitrage.

We find that shocks to funding liquidity and to variables that proxy for the intensity of arbitrage activity have an economically and statistically significant impact on systematic market efficiency. In particular, a negative shock to the TED spread and a positive shock to hedge fund assets under management or to algorithmic trading positively affect the systematic component of market efficiency, both contemporaneously and in subsequent months. These results indicate that, consistent with recent theories, funding liquidity and the intensity of arbitrage activity are important factors that help us understand the driving forces of systematic variation in stock market efficiency. Furthermore, we document that the effect of hedge fund assets under management on market efficiency is greater for high turnover stocks than for low turnover stocks, while the effect of the TED spread is more pronounced for low turnover stocks.

To our knowledge, no previous work studies the degree and determinants of systematic variation in market efficiency measures for individual stocks. We view our analysis as relevant for at least two reasons. First, we show that stock market efficiency, rather than being a static concept, exhibits significant variation over time, and that different efficiency measures co-move across individual stocks as well as with each other. Second, we note that while prior work has studied the link between funding liquidity and market liquidity (e.g., Brunnermeier and Pedersen, 2009; Hameed, Kang, and Viswanathan, 2010), and between funding liquidity and *specific* arbitrage strategies in convertible bonds, mergers, covered interest parity, credit default swaps, and closed-end funds (e.g., Mitchell, Pedersen, and Pulvino, 2007; Gârleanu and Pedersen, 2011; Mancini-Griffoli and Ranaldo, 2011; Mitchell and Pulvino, 2012), our study demonstrates a connection between funding liquidity and the systematic component of commonly accepted efficiency measures for equities. Our results suggest that policy attempts to increase funding liquidity may not only have a direct impact on trading costs, but also on the systematic degree of stock price efficiency. Further, our results are complementary to Pasquariello's (2014) important study of fluctuations in financial market dislocations, which are constructed as an average of violations of arbitrage parities across stock, foreign exchange, and money markets. Our analysis instead focuses on individual stocks (and stock options) and indicates that the price efficiency of individual stocks fluctuates over time in a systematic way.

1. Efficiency measures

Market efficiency is a central concept in finance, and academic research has a longstanding interest in measuring the extent to which financial markets or individual securities exhibit efficient price formation. A number of distinct efficiency measures have been developed in the literature. Some of these measures are designed to capture the extent to which stock prices deviate from a random walk (e.g., return predictability, variance ratios), while others aim to measure pricing errors relative to the efficient market benchmark (e.g., Hasbrouck, 1993) or violations of the law of one price across different markets (e.g., put-call parity deviations). All of these measures have been used in different lines of research.

Our purpose is to analyze the extent to which these different efficiency measures co-move over time, both across individual stocks and with each other, and to examine the determinants of any systematic variation in efficiency across stocks and measures. We therefore focus on efficiency measures that can be estimated at the stock-level and at a relatively high frequency. Our search of the literature identifies four different measures that are widely used and that can be estimated daily for a large cross-section of stocks based on high-frequency data: intraday return predictability, variance ratios, Hasbrouck's (1993) pricing errors, and put-call parity deviations. Further, in our analysis of co-movement across aggregate efficiency measures, we include the market-wide cross-sectional return predictability measure of Heston, Korajczyk, and Sadka (2010). We now explain how we estimate these measures (Section 1.1. through Section 1.5) and discuss the relation between the efficiency measures and market liquidity (Section 1.6).

1.1 Intraday return predictability

Our first measure is based on the intraday predictability of individual stock returns from past order flow or past returns. Several papers, including Hasbrouck and Ho (1987), Chan and Fong (2000), Chordia, Roll, and Subrahmanyam (2005), and Boehmer and Wu (2007), explore and provide evidence of such return predictability, which we use as an inverse indicator of market efficiency. Chordia, Roll, and Subrahmanyam (2005) suggest that such predictability arises from dealers' risk aversion, which delays the accommodation of autocorrelated order imbalances. Their evidence suggests that trading by astute arbitrageurs removes all return predictability over intervals of five minutes or more, but some predictability remains at shorter horizons.

In line with these prior studies, we estimate the intraday return predictability of each individual stock for each day in the sample based on regressions of stock returns over short intervals within the day on order imbalance (dollar volume of buyer- minus seller-initiated trades) in the previous interval. Chordia, Roll, and Subrahmanyam (2005) show that prices cease to be predictable from order flow in 30 minutes or less in 1996, and in around five minutes in 2002. Since our sample period lasts till 2010, it is judicious to use intervals shorter than five minutes to still capture meaningful predictability in the later part of the sample period. In light of this consideration, we estimate predictability based on intraday returns and order imbalances measured over one-minute intervals (with a robustness check based on two-minute intervals).

We estimate the extent of short-horizon return predictability from order flow for each stock i and day d in the sample as the slope coefficient from the following regression, using intraday data aggregated over one-minute intervals:

$$R_{i,d,t} = a_{i,d} + b_{i,d} OIB_{i,d,t-1} + \epsilon_{i,d,t}, \tag{1}$$

where $R_{i,d,t}$ is the return of stock i in one-minute interval t on day d based on the mid-quote associated with the last trade to the mid-quote of the first trade in the interval (we use mid-quote returns to avoid the bid-ask bounce), and $OIB_{i,d,t-1}$ is the order imbalance for the same stock and day in the previous interval $t-1$, computed as the difference between the total dollar volume of trades initiated by buyers and sellers ($OIB\$$). A smaller slope coefficient b from the regression in Eq. (1) indicates greater efficiency. We refer to the efficiency measure based on this regression specification as *OIB predictability*.

To assess the robustness of our results to changes in the specification of the predictability regressions, we also estimate four alternative return predictability measures, each named after the single feature that distinguishes it from the *OIB predictability* measure. The *allquotes* measure is based on returns computed using all quotes within each interval rather than only using quotes associated with trades; the *2minutes* measure is based on two-minute instead of one-minute intervals; and the *oib#* measure is based on order imbalance expressed in number of trades rather than dollars. We also present and discuss the results using the slope coefficient from regressions of one-minute returns on their one-minute lagged counterparts, instead of past order flows, and label this the *autocorrelation* measure. We discard stock-days with fewer than 20 observations for each of these measures. In our analyses of co-movement in market efficiency, we use a comprehensive *Predictability* measure that is constructed as the first principal component across the five alternative return predictability measures (more details are provided below).

1.2 Variance ratios

The second stock-level efficiency measure we consider is a daily variance ratio that examines how closely the price of individual stocks adheres to a random walk benchmark; this measure is in line with, among others, Bessembinder (2003). The stock-level *Variance ratio* measure is defined as $|1 - 30 \times Var(1\text{min})/Var(30\text{min})|$, where $Var(1\text{min})$ is the return variance estimated from one-minute mid-quote returns within a day and $Var(30\text{min})$ is the return variance estimated from 30-minute mid-quote returns within a day. Variance ratios are computed from mid-quote returns and do not utilize traded prices, mitigating the problem of non-synchronous trading. Since estimates of daily variance ratios of individual stocks can be noisy (Andersen, Bollerslev, and Das, 2001), we follow Lo and MacKinlay (1989; see their equation (5)) and Charles and Darné (2009) and estimate daily variance ratios based on overlapping intraday returns. Since expected returns over such short intervals are very close to zero, we set expected returns to zero in the computation of the variances. We discard stock-days with fewer than 20 non-zero one-minute returns. The *Variance ratio* measure tends to unity as serial dependence in asset returns tends to zero; therefore, it measures how closely the price adheres to a random walk.

1.3 Hasbrouck pricing errors

As a third daily, stock-level efficiency measure, we estimate Hasbrouck's (1993) pricing errors based on intraday trades and quotes. Hasbrouck proposes a method to decompose stock prices into random walk and stationary components. He refers to the stationary component (the difference between the efficient price and the actual price) as the pricing error, which he argues is a natural measure for price inefficiency. We follow Hasbrouck and estimate vector autoregression (VAR) models to estimate these components. As in Boehmer and Kelley (2009), we estimate a five-lag VAR model based on intraday data for each stock-day with at least one hundred trades. The endogenous variables of the model are: (i) the logarithmic price return, from quote midpoints associated with trades,[5] (ii) a trade sign indicator, (iii) the signed volume (that is, the sign of the trade times the number of shares traded), and (iv) the sign of the trade times the square root of the number of shares traded. We sign all trades with trade prices above the prevailing quote midpoint as buyer-initiated, and seller-initiated if they are below the quote midpoint. If the trade occurred at the prevailing quote midpoint we set the sign of the trade to zero (following Hasbrouck, 1993). As in Hasbrouck (1993), we set all lagged variables at the beginning of each day to zero. We obtain the pricing error of each trade in a stock on a

[5]Using mid-quote returns avoids the bid-ask bounce, but using returns from actual trade prices does not alter the main results.

given day from the vector moving average representation of the VAR system (Beveridge and Nelson, 1981) using Eq. (13) in Hasbrouck (1993).

Prior studies use the standard deviation of the intraday pricing errors as an inverse measure of informational efficiency. However, for our purpose, we are more interested in the magnitude of the pricing error rather than in its intraday variation. We thus take the maximum of the absolute pricing errors of the trades in a stock on a given day as an inverse measure of the price efficiency for that stock on that day and label it the *Hasbrouck* measure. Since daily stock-level estimates of the maximum intraday pricing error exhibit several large outliers, we use the logarithmic transformation of *Hasbrouck* to mitigate their influence.

1.4 Put-call parity deviations

Our fourth daily proxy for the price efficiency of individual stocks is a law of one price measure derived from options markets. The use of this measure enhances our understanding of co-movement in market efficiency by extending the notion of efficiency to derivatives markets for individual stocks. This *Put-call parity* measure is estimated using the OptionMetrics database as the absolute difference between the implied volatilities of a call and a put option of the same series (i.e., pairs of options on the same underlying stock with the same strike price and the same expiration date).[6] We use end-of-day quotes from all option series with positive implied volatilities that expire in two weeks to one year and that have a strike-to-spot ratio between 0.95 and 1.05. This ensures that our estimates of put-call parity deviations are based on what are typically the most liquid options (following Pan, 2002). When more than one option pair satisfies these conditions

[6]This measure is also used in Cremers and Weinbaum (2010). These authors note that while, strictly speaking, put-call parity does not hold as an equality for the American options on individual stocks, a lower discrepancy in implied volatilities from binomial models nonetheless is indicative of more efficient options and stock markets.

for a given stock-day, we take the average of the absolute differences between the implied volatilities of the call and the put option across all option pairs.

1.5 Cross-sectional intraday return predictability (HKS)

In our analysis of co-movement in efficiency across measures, we also include a monthly, market-wide efficiency measure based on Heston, Korajczyk, and Sadka (2010). These authors document a remarkable pattern of cross-sectional predictability of intraday returns: stocks with a relatively high 30-minute return at a particular time during the trading day tend to also have a relatively high return at the same time on the next trading day. They argue that the combination of autocorrelated institutional investment flows and optimal trading strategies gives rise to predictable patterns in trading that are not fully anticipated by the market. Following their approach, we divide the 6.5-hour trading day into thirteen 30-minute intervals and run cross-sectional regressions of 30-minute stock returns on returns over the same interval on the previous day. In line with Heston, Korajczyk, and Sadka (2010), we take the slope coefficient in these regressions (averaged over all intervals within a month) as a monthly, market-wide measure of efficiency and refer to it as the *HKS* measure.

1.6 Relation between efficiency measures and market liquidity

One issue that arises in all analyses of (common) time-variation in the different market efficiency measures included in this paper is how these measures are related to market liquidity. Characterizing the relation between efficiency and liquidity is not straightforward, since the causality can run either way, since competing hypotheses predict opposite-sign relations, and since the relation may depend on the specific efficiency measure used.

We first note that illiquidity does not necessarily imply return predictability from order flow or past returns (*Predictability* and *Variance ratio* measures) or pricing errors relative to efficient prices (the *Hasbrouck* measure). In Glosten and Milgrom (1985) and Kyle (1985), even though markets are illiquid, price changes are serially uncorrelated because market makers are risk-neutral. However, there are alternative channels that could give rise to inefficiencies. To discuss these channels, it may be useful to consider the following taxonomy of agents: traders who demand immediacy for liquidity or informational needs, liquidity providers (both designated market makers or specialists and de facto market makers such as algorithmic traders), and outside arbitrageurs who exploit deviations from efficient prices.

In inventory-based models such as Stoll (1978), efficiency can be compromised if market makers have capital constraints or limited risk-bearing capacity, inhibiting their ability to prevent prices moving away from fundamentals as a result of demand or supply shocks from liquidity traders. Alternatively, such shocks can also result in inefficiencies when market makers are risk-neutral but face cognitive limitations and thus might misreact to the information content of the order flow (Barberis, Shleifer, and Vishny, 1998). Inefficiencies resulting from these channels may be reflected in all five efficiency measures used in this paper. The *Predictability* and the *Variance ratio* measures are designed to pick up return predictability from order flow or return autocorrelations resulting from either of these channels. *Hasbrouck* pricing errors may also stem from the inventory-based channel, since they can be viewed as the result of, among others, "inventory control" and "the transient component of the price response to a block trade" (Hasbrouck, 1993, pp. 193-194.) And although Hasbrouck (1983) does not explicitly consider market makers' potential cognitive limitations, these too could arguably lead to price deviations from the efficient market benchmark as reflected in the *Hasbrouck* measure. The *Put-call parity* measure may indicate greater deviations from the law of one price when price pressures temporarily move prices in either the stock or the options market. Further, *HKS* efficiency may deteriorate when market makers fail to counteract predictable patterns in the cross-section of intraday stock returns because of either inventory concerns and/or cognitive limitations.

In a third channel, efficiency might be challenged as a result of informational differences when price adjustments to information in asset pairs with common fundamentals occur asynchronously due to, for example, "lags in the transmission and interpretation of prices" (Kumar and Seppi, 1994), "poor intermarket information linkages" (Domowitz, Glen, and Madhavan, 1998), or "stale quotes" (Foucault, Kozhan, and Tham, 2015). These three papers study such inefficiencies in the context of, respectively, index arbitrage, cross-listings arbitrage, and triangular arbitrage in foreign exchange markets. The informational differences channel may thus be most directly relevant for law of one price deviations, and thus for the *Put-call parity* measure among the five efficiency measures considered in this paper.

Based on this taxonomy, we can derive three alternative rationales for the relation between efficiency and liquidity. The first rationale considers the effect of liquidity on efficiency. In all three channels through which inefficiencies can arise, outside arbitrageurs who monitor the market may detect temporary deviations from efficient prices and may submit arbitrage orders to exploit such inefficiencies. To the extent that they use market orders (or marketable limit orders) to ensure speedy execution in active markets in which inefficiencies might be short-lived, they will be discouraged to do so when the bid-ask spread, a measure of illiquidity, is large (Chordia, Roll, and Subrahmanyam, 2008). Hence, in line with the limits to arbitrage literature, illiquidity is a potentially important friction that hampers the ability of arbitrageurs to restore market efficiency. Since each efficiency measure used in this paper is linked to arbitrage, this hypothesis thus predicts a positive relation between market liquidity and all five efficiency measures.

The second rationale considers the relation between efficiency and liquidity in the inventory-based channel. In this channel, outside arbitrageurs may effectively complement the capital and/or risk-bearing capacity of the market making sector by acting as de facto liquidity providers (Holden, 1995; Gromb and Vayanos, 2010; Nagel, 2012), for example by submitting limit orders. In this case, arbitrage activity enhances efficiency and liquidity in chorus. Since the inventory-based channel may give rise to inefficiencies that are reflected in each of the efficiency measures in this paper, this hypothesis also predicts a positive relation between efficiency and liquidity.

The third rationale considers the relation between efficiency and liquidity in the cognitive limitations and informational differences channels. In these channels, arbitrage activity may exacerbate the market makers' adverse selection problem, since arbitrageurs exploit their cognitive limitations or trade on informational differences across markets. This rationale thus predicts that arbitrage could decrease market liquidity, suggesting a negative relation between efficiency and liquidity. As discussed above, the cognitive limitations channel may be relevant for all five efficiency measures, while the informational differences channel is most pertinent for the *Put-call parity* measure.

To summarize, while liquidity and efficiency are distinct concepts, there are several reasons to expect a relation between these concepts, and they may apply to a lesser or greater degree to the different efficiency measures we consider. In this paper, while we recognize the link between efficiency and liquidity, we desist from discerning between the different explanations for this link. But, if our efficiency measures overlap considerably with illiquidity, our analysis of co-movement in efficiency across stocks might be perceived as a reiteration of the extensive literature on co-movement in liquidity (e.g., Chordia, Roll, and Subrahmanyam, 2000; Hasbrouck and Seppi, 2000; Huberman and Halka, 2001). Therefore, we first orthogonalize each of the four daily efficiency measures at the stock-level with respect to a measure of that stock's illiquidity.

What illiquidity measure is most appropriate? The recommended illiquidity proxy in Hasbrouck (2009) and Goyenko, Holden, and Trzcinka (2009) is the monthly Amihud (2002) measure. However, we perform analyses at both the daily and monthly frequencies, and the daily Amihud (2002) measure tends to be quite noisy. Thus, for our daily analyses, we orthogonalize our efficiency measures with respect to the daily proportional quoted bid-ask spread or *PQSPR* (computed as the time-weighted average over the trading day of the bid-ask spread scaled by the quote midpoint). We then run our analyses of co-movement in efficiency across stocks in Section 3.1 using the orthogonalized daily, stock-level efficiency measures. We obtain similar results when we orthogonalize with respect to the daily proportional effective spread or *PESPR* (computed as the average across all trades on a day of two times the absolute difference between the transaction price and the quote midpoint, scaled by the quote midpoint) and slightly stronger results when we do not orthogonalize at all. We also obtain similar results when we orthogonalize with respect to the daily Amihud (2002) illiquidity proxy (computed as the daily ratio of the absolute stock return to dollar trading volume, cross-sectionally winsorized at the 99.5% each day to mitigate the influence of outliers).

For the analyses of time-variation in the monthly, market-wide efficiency measures in Section 3.2 and Section 4, we orthogonalize the monthly, stock-level *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity* measures with respect to the monthly, stock-level *Amihud* measure (computed as the average across all trading days within the month of the daily ratio of the absolute stock return to dollar trading volume, cross-sectionally winsorized at the 99.5% level each month) before aggregating the stock-level efficiency measures to the market-level by value-weighting across stocks. In the same vein, we orthogonalize the monthly, market-wide *HKS* measure with respect to the monthly, market-wide *Amihud* measure. We choose to report the results based on monthly efficiency measures orthogonalized with respect to the monthly *Amihud* measure. However, our main results are not materially affected when we orthogonalize the monthly efficiency measures with respect to the monthly *PQSPR* or *PESPR* (each computed as the average across all trading days within the month of the daily *PQSPR* and *PESPR* measures).

2. Sample and efficiency estimates

This section discusses the data sources and screens (Section 2.1) and presents the estimates of the daily, stock-level efficiency measures (Section 2.2).

2.1 Data and sample

To estimate the five efficiency measures, we obtain data on all trades and quotes as well as their respective sizes for individual U.S. stocks from the Thomson Reuters Tick History (TRTH) database, which contains global tick-by-tick trade and quote data across asset classes.[7] Our data start in March 1996, which is the earliest month available in the TRTH database. Our sample consists of all NYSE stocks that were traded at any time during our sample period from March 1996 to December 2010 and that survive our data screens. We include only NYSE stocks to prevent issues with differences in trading volume definitions across NYSE and Nasdaq, see, e.g., Gao and Ritter (2010). We use trades and (national best bid and offer or NBBO) quotes on all U.S. exchanges on which these NYSE stocks are traded. We apply a variety of filters to the data that are described in the online appendix.[8] Our final sample includes 2,157 NYSE stocks.

To estimate the predictability regressions in Eq. (1), we require at least one signed trade in both the interval over which we calculate the return as well as the previous

[7]To verify that our results do not depend on using TRTH instead of NYSE's Trade and Quote (TAQ) database, we compare the results based on TRTH to those based on TAQ for all 2,023 NYSE-listed common stocks that were traded at any time over the period 1996-2000 and find that they are very similar.

[8]This appendix also presents results from the robustness checks mentioned within the paper.

interval. We discard stock-days for which we have fewer than 20 one-minute intervals with valid data on the stock return within that interval and on the order imbalance or return in the preceding interval (in total 756,051 stock-day observations), and days for which TRTH reports a data gap that overlaps with the continuous trading session (in total 56 days). Our data filters allow us to estimate Eq. (1) for on average around 1,700 days over the period 1996-2010 for around 1,900 stocks in our sample (depending on the predictability measure). We are able to use 14,253,093,209 transactions, signed by the Lee and Ready (1991) method, in our analyses.[9]

Table 1 presents summary statistics of the return and order imbalance variables that serve as inputs to our predictability regressions. For these variables, the table reports cross-sectional summary statistics (the mean, standard deviation, as well as the median and the 25th and 75th percentiles) of the stock-by-stock time-series averages. The average number of trades per day is around 2,000. The average daily dollar trading volume is 0.025 or US$25m. The median one-minute mid-quote return is equal to -0.001 basis point, which corresponds to -0.4 basis points per day. The negative median return is likely driven by the fact that intraday returns tend to be lower than overnight returns (Berkman, Koch, Tuttle, and Zhang, 2012, report negative mean and median open-to-close returns for a sample of 3,000 U.S. stocks over 1996-2008). There is a slight positive average order imbalance over the one-minute intervals in our sample.

2.2 Daily, stock-level efficiency estimates

Panel A of Table 2 presents the results of the daily return predictability regressions estimated based on intraday data. As described in Section 1.1, the baseline predictability

[9] The Lee/Ready algorithm classifies a trade as buyer- (seller-)initiated if it is closer to the ask (bid) of the prevailing quote. If the trade is exactly at the midpoint of the quote, the trade is classified as buyer- (seller-)initiated if the last price change prior to the trade is positive (negative). Lee and Radhakrishna (2000) and Odders-White (2000) indicate that the Lee/Ready algorithm is quite accurate for NYSE stocks, suggesting that assignment errors should have minimal impact on the results.

measure (*OIB predictability*) is obtained from regressions of one-minute mid-quote returns (computed using quotes associated with trades) on lagged dollar order imbalance. For robustness, we also estimate four alternative predictability measures: *allquotes*, *2minutes*, *oib#*, and *autocorrelation*.

Consistent with prior research, Panel A of Table 2 shows that order imbalance positively predicts future returns over short intervals. The average coefficient on lagged order imbalance across the approximately 3.2 million stock-day regressions ranges from 0.947 for the *oib#* measure to 6.169 for the *2minutes* measure. The return autocorrelation coefficient is also positive at 0.024. The first number below the average coefficient in each column ("t-stat avg") is the average t-statistic across all stock-day regressions. Although for all measures except perhaps one (*oib#*), the simple average t-statistic does not exceed critical values associated with conventional confidence levels, the t-statistics of the individual stock-day regressions can be based on as few as 20 intraday observations. The second number below the average coefficient in each column ("NW t-stat avg"), is the Newey-West (1994) t-statistic computed based on the time-series of daily coefficient estimates of individual stocks, which is then averaged across stocks. These statistics are highly significant for all five predictability regressions reported in Panel A of Table 2 and indicate that intraday returns exhibit significant predictability from lagged order imbalance or returns.

Panel A of Table 2 also shows that a large fraction (around 60-90%, depending on the predictability measure) of the coefficients on lagged order imbalance and on lagged returns in the individual stock-day predictability regressions are positive, and that 30-60% of these coefficients are significant on an individual basis. The average R^2 of the regressions ranges from 1.7% for *allquotes* to 3.5% for *oib#*. Although these R^2's are modest, we note that predicting stock returns is challenging and that the results are in line with prior work on intraday return predictability (e.g., Chordia, Roll, and Subrahmanyam, 2005).[10]

Overall, Panel A of Table 2 provides evidence of significant intraday return predictability in our sample of all NYSE stocks over 1996-2010. The results also indicate that the degree of predictability is robust across various specifications of the predictability regressions.

Panel B of Table 2 presents cross-sectional summary statistics of the stock-by-stock time-series averages of the five different return predictability measures as well as the other three stock-level efficiency measures (*Variance ratio*, *Hasbrouck*, and *Put-call parity*). This panel is based on the sample of stocks for which each efficiency measure could be estimated for at least 15 days over the sample period.[11]

To compress the five return predictability measures in Table 2 into a single measure to be used in the remainder of the paper, for each stock we take the first principal component of the daily time-series of slope coefficients of the five different predictability regressions in Panel A and label it the *Predictability* measure. On average, this first principal component explains more than 50% of the total variation in the five predictability measures for individual stocks. The loadings on the first principal component almost always have the same sign for all five predictability measures, with the exception of 156 stocks (out of the 1,827 stocks for which we can estimate all five predictability regressions). We obtain similar results when we drop these 156 stocks from the sample. The average

[10]We also estimate an intraday predictability measure based on regressions that include lagged order imbalance in dollars and in trades as well as lagged returns simultaneously, and find considerably stronger return predictability based on all three variables.

[11]We note that although the literature on intraday return predictability (e.g., Boehmer and Wu, 2007; Andrade, Chang, and Seasholes, 2008) presents overwhelming evidence that intraday returns are predicted positively by lagged order flow and lagged returns, and although the vast majority of the estimated slope coefficients in Panel A of Table 2 are positive, (large) negative slope coefficients in Eq. (1) for a particular stock-day could arguably also be interpreted as evidence of inefficiency. However, our main results are not materially affected when we take the absolute slope coefficient from Eq. (1) as our stock-day return predictability measure or when we set negative stock-day coefficients to zero.

loadings of the first principal component on the underlying predictability measures are 0.54 for *OIB predictability*, 0.48 for *allquotes*, 0.42 for *2minutes*, 0.42 for *oib#*, and 0.14 for *autocorrelation*, which indicates that the resulting *Predictability* measure is fairly representative of the various individual intraday return predictability measures.

The mean and median absolute deviations of the *Variance ratio* from unity reported in Panel B of Table 2 are equal to 0.87 and 0.76, respectively. These numbers are somewhat higher than the mean of 0.53 reported by Boehmer and Kelley (2009, see their Table 1), but that number is based 1-to-20 days variance ratios (instead of 1-to-30 minutes variance ratios as in our paper) and based on a sample of NYSE stocks that is about half the size of our sample and likely tilted towards large and liquid stocks that may be more efficiently priced.

The mean (median) value of the *Hasbrouck* measure is 39 (24) basis points. These numbers align well with the mean pricing error of 26 basis points reported by Hasbrouck (1993) for a representative sample of 175 NYSE stocks in 1989. We would expect pricing errors to be lower in our more recent sample, but we report the maximum rather than the mean pricing error.

We are able to estimate the *Put-call parity* measure for 1,535 of the 2,157 stocks in our sample, for on average 1,448 days over our sample period 1996-2010. The mean absolute put-call parity deviation (expressed in terms of implied volatility) across stock-days in the sample is 2.58%, with an interquartile range of 1.60%. These values closely correspond to the put-call deviation estimates provided by Cremers and Weinbaum (2010) for a similarly-sized sample of U.S. stocks over 1996-2005. Panel A of their Table 1 shows an average put-call parity deviation of −0.978%, but this is an aggregation of positive and negative deviations. Taking the average of the absolute values of the percentiles of the distribution of their put-call parity deviation estimates reported in Panel B of their Table 1 yields an approximate average absolute deviation of 2.3% for their sample.

All of the stock-level efficiency measures in Panel B of Table 2 show large cross-sectional standard deviations and interquartile ranges, demonstrating that the degree of price efficiency varies considerably across individual stocks. In addition, there is substantial time-variation in the different stock-level efficiency measures. As an illustration, the market-wide (equally-weighted average) R^2 of the *OIB predictability* regressions is 6.44% in 1996 but only 1.29% in 2010. The average across stocks of the stock-by-stock time-series standard deviation of the *Variance ratio*, *Hasbrouck*, and *Put-call parity* measures is 1.08, 0.51, and 2.73, respectively (not tabulated to conserve space). These average standard deviations are all large relative to the average across stocks of the stock-by-stock time-series averages of these measures of 0.87, 0.39, and 2.58, respectively (from Panel B of Table 2).

Panel C of Table 2 presents average cross-sectional Spearman rank correlations across the monthly, stock-level *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity* measures (Pearson correlations are similar). We construct these monthly, stock-level efficiency measures by averaging the corresponding daily, stock-level measures across days within the month to mitigate the noise inherent in the individual stock-day efficiency estimates. Most of the correlations are both economically and statistically significant, which indicates that although the degree of price efficiency varies considerably across stocks, the different efficiency measures tend to provide a similar indication of the relative degree of price efficiency of individual stocks. The exception is the cross-sectional correlation between the *Predictability* and *Variance ratio* measures, which is economically small and statistically indistinguishable from zero.

3. Co-movement in efficiency measures

We now examine whether there is co-movement in different market efficiency measures across stocks (Section 3.1) and across measures (Section 3.2).

3.1 Co-movement in efficiency across stocks

To estimate the degree of co-movement in efficiency across stocks, we run time-series regressions of the efficiency of individual stocks on contemporaneous, lead, and lagged market-wide efficiency. Specifically, we estimate the degree of co-movement in efficiency for each stock i over the whole sample period in the following regression:

$$Eff_{i,d} = \alpha_i + \beta_i MktEff_{i,d} + \gamma_i MktEff_{i,d-1} + \delta_i MktEff_{i,d+1} + \eta_{i,d}, \qquad (2)$$

where $Eff_{i,d}$ is the efficiency of stock i on day d, and $MktEff_{i,d}$ is the market-wide efficiency (defined as the value-weighted average efficiency across all stocks in our sample excluding stock i). We estimate Eq. (2) for each stock with at least 15 daily observations over the whole sample period, based on daily estimates of our four stock-level efficiency measures: *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity*.

As discussed in Section 1.6, our analysis of co-movement in efficiency across stocks is based on stock-level efficiency measures that have been orthogonalized with respect to stock-level liquidity. In particular, we run regressions as in Eq. (2) of stock-level efficiency orthogonalized with respect to liquidity on contemporaneous, lead, and lagged orthogonalized market efficiency (defined as the value-weighted average efficiency, orthogonalized with respect to liquidity, across all stocks in our sample, excluding stock i). In robustness tests, we also estimate the co-movement regressions in Eq. (2) based on efficiency changes orthogonalized with respect to liquidity changes rather than based on efficiency levels orthogonalized with respect to liquidity levels, and based on contemporaneous market efficiency as the only independent variable (that is, no lead and lagged market-wide efficiency), and obtain similar results. We also obtain similar results when we compute market-wide efficiency as the equally-weighted (instead of the value-weighted) average efficiency across all stocks in our sample, excluding stock i.

Table 3 presents the results of our regressions to estimate co-movement in each of the four efficiency measures across individual stocks. The table reports the average regression coefficients across all co-movement regressions estimated by stock for each efficiency measure. The number of stocks for which we can estimate Eq. (2) varies from 1,505 for the *Put-call parity* measure to 2,041 for the *Variance ratio* measure.

The table reveals evidence of significant co-movement in efficiency across stocks. The average coefficient on contemporaneous market-wide efficiency across the regressions estimated for individual stock is positive and economically substantial for all efficiency measures, ranging from 0.483 for the *Predictability* measure to 1.528 for the *Hasbrouck* measure. The average t-statistic of these coefficients (the first number below the average coefficient on contemporaneous market efficiency in each column) is highly significant, indicating that, on average, the estimated coefficient on contemporaneous market efficiency is at least four standard errors away from zero.

This conclusion is confirmed by the second number below the average coefficient on contemporaneous market efficiency in each column, which is the t-statistic computed from the cross-sectional distribution of estimated coefficients of all stocks ("CS t-stat avg"). These t-statistics are corrected for cross-correlations in the residuals of the individual regressions using the method outlined by Chordia, Roll, and Subrahmanyam (2000, 2008). In line with their recommendation, for each column in Table 3, we compute the average pairwise correlation (ρ) between the residuals across all N regressions and then multiply the standard errors by $[1 + (N - 1)\rho]^{1/2}$. The resulting adjusted t-statistics also indicate statistically significant co-movement in efficiency across individual stocks for all four efficiency measures.

For each of the four efficiency measures, a clear majority (at least 87% and up to 92%) of the individual coefficients on contemporaneous market-wide efficiency are positive. At least 66% (*Hasbrouck*) and up to 72% (*Variance ratio*) of the coefficients are positive and significant on an individual basis. There is little evidence that the lead and lagged market-wide efficiency are important in explaining time-variation in the efficiency of individual stocks.

In the spirit of Morck, Yeung, and Yu (2000), the R^2 from the co-movement regressions in Eq. (2) can be interpreted as an alternative way to gauge the degree of efficiency co-movement across stocks. The average (adjusted) R^2's of the co-movement regressions in Table 3 range from 3.2% (2.3%) for *Variance ratio* to 8.8% (8.0%) for *Put-call parity.* However, the degree of co-movement in efficiency uncovered in Table 3 may be mitigated because we restrict the coefficients to be the same over the whole sample period. Indeed, when we estimate Eq. (2) for each stock over each month (instead of over the whole sample period), we obtain average R^2's of around 20-22%, which are comparable to the R^2 of around 23% reported by Karolyi, Lee, and van Dijk (2012) for similar monthly regressions to estimate co-movement in liquidity for NYSE stocks over the period 1995-2009. We also note that considering portfolios of stocks might alleviate estimation noise and expose a stronger image of co-movement. Accordingly, we revisit the co-movement analysis in Table 3 by running regressions of the value-weighted average efficiency of the stocks in each of the five industries defined on the website of Ken French on contemporaneous, lead, and lagged market-wide efficiency (computed as the value-weighted average efficiency across the stocks not in the subject industry). In the online appendix, we show that the industry-level R^2's of the co-movement regressions in Eq. (2) range from 20% to 80%

and are thus considerably greater than the individual stock-level R^2's reported in Table 3.

Overall, Table 3 presents evidence of economically and statistically significant co-movement in efficiency across stocks, suggesting that common economic forces may drive variation in the price efficiency of individual securities.

3.2 Co-movement in efficiency across measures

We now turn to the question whether there is evidence of common variation in aggregate efficiency across different efficiency measures. As discussed in Section 1.6, we first orthogonalize the monthly, stock-level *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity* measures from Panel C of Table 2 with respect to monthly, stock-level liquidity. Subsequently, we compute the value-weighted average efficiency (orthogonalized with respect to liquidity) across individual stocks each month, separately for each efficiency measure. This procedure yields four different monthly, market-wide efficiency measures. We include the *HKS* measure (orthogonalized with respect to the value-weighted average liquidity across individual stocks) as a fifth monthly, market-wide efficiency measure in our analyses.[12]

Figure 1 shows the monthly time-variation in the five market-wide efficiency measures, each of which represents an inverse measure of informational efficiency, so that lower values indicate greater efficiency. All measures show considerable time-variation over our sample period. Visual inspection suggests that there are some clear common patterns across several of the measures, but also some idiosyncratic fluctuations. For

[12]Our data are based on all national best bid and offer (NBBO) quotes on all U.S. exchanges, while the regressions in Heston, Korajczyk, and Sadka (2010) are based on NYSE data only. Also, we use mid-quote returns (instead of return from trade prices) to avoid bid-ask bounce. Nonetheless, we are able to closely match their results. In particular, they report an average slope coefficient on 30-minute returns that are lagged 13 intervals (or one full trading day) of 1.19 (their Table I), while we obtain an estimate of 1.15 over their sample period 2001-2005.

example, *Predictability*, *Hasbrouck*, and *Put-call parity* all show a long-run downward sloping trend, which is particularly pronounced over the period 2000-2008 and indicates a gradual overall improvement in efficiency over our sample period. *Hasbrouck*, *Put-call parity*, and *HKS* each indicate a marked deterioration in efficiency following the collapse of Lehman Brothers in September 2008. *Predictability*, *Hasbrouck*, and *Put-call parity* show a similar drop in market efficiency around the LTCM crisis in August 1998. On the other hand, time-variation in *HKS* and most notably *Variance ratio* seems largely independent of time-variation in the other measures.

In Table 4, we assess the extent of co-movement across the five different monthly, market-wide efficiency measures by presenting Spearman rank correlations across all of the measures (Pearson correlations are similar). The main takeaway from the table is that correlations across the efficiency measures tend to be economically substantial and statistically significant, with the notable exception of those involving the *Variance ratio* measure. For example, the *Hasbrouck* measure has a correlation of 59.75% with *Predictability*, 42.07% with *Put-call parity*, and 14.87% with *HKS*. The correlation between *Put-call parity* and *Predictability* is 18.83%. All of these correlations are statistically significant. Correlations between the *HKS* measure and the *Predictability* and *Put-call parity* measures are weak.[13]

The results in Table 4 point to significant common variation across four of the five market-wide efficiency measures. Remarkably, variation in the *Variance ratio* measure is negatively correlated with variation in the other efficiency measures, perhaps underlining the concerns about stock-level variance ratios expressed by Andersen, Bollerslev, and Das

[13]We note that the correlations in Table 4 may in part be driven by a common time trend. We acknowledge that correlations based on trending variables could to some degree be spurious. At the same time, we are interested in the extent to which different efficiency measures reflect any trends in aggregate efficiency in a similar way. In the same vein, Goyenko, Holden, and Trzcinka (2009) do not detrend the underlying variables in their analysis of the extent of co-variation across high- and low-frequency measures of market liquidity.

(2001). The correlations of *Variance ratio* with *Predictability*, *Hasbrouck*, and *Put-call parity* are all around -35% and statistically significant, while the correlation with *HKS* is also negative but insignificant.

We conclude that there is an important common component to the variation in various market-wide efficiency measures, but correlations are less than perfect (most notably for variance ratios). Given these imperfect correlations, it seems reasonable to rely upon more than a single measure of market efficiency in empirical research.

4. What drives systematic variation in market efficiency?

Our results in Section 3 indicate that market efficiency measures exhibit significant co-movement across individual stocks and with each other. These findings suggest that there is a significant systematic component to these market efficiency measures. We now turn to an analysis of what economic forces drive time-variation in the systematic market efficiency component. We first develop hypotheses and discuss the variable definitions (Section 4.1) and then present our measure of systematic market efficiency (Section 4.2), our methodology (Section 4.3), and the empirical results (Section 4.4).

4.1 Hypotheses development and variable definitions

Since market efficiency is enforced in part by way of arbitrage, factors that affect the efficacy of arbitrage could induce systematic variation in efficiency. Recent theoretical and empirical research suggests that changes in the availability of arbitrage capital are an important source of variation in arbitrage efficacy. For example, in the theoretical models of Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009), Gârleanu and Pedersen (2011), and Liu and Mello (2011), arbitrageurs are capital constrained and, as a result, may not be able to eliminate deviations from efficient pricing. In several of these models, adverse shocks to funding liquidity force arbitrageurs to terminate their positions, which can lead to greater deviations from efficient pricing and potentially trigger feedback loops that cause inefficiencies and funding constraints to worsen in lock-step. Empirical studies by Mitchell, Pedersen, and Pulvino (2007), Mancini-Griffoli and Ranaldo (2011), and Mitchell and Pulvino (2012) document that funding liquidity is subject to significant fluctuations, and that improvements in funding liquidity are associated with increased arbitrage efficacy. We thus hypothesize that improved funding liquidity enhances systematic market efficiency.

Our key proxy for the overall level of funding liquidity is the *TED spread*, which is the difference between the three-month LIBOR and the three-month T-bill rate from the FRED database of the Federal Reserve Bank of St. Louis and is a widely used indicator of funding liquidity (Brunnermeier, Nagel, and Pedersen, 2008; Brunnermeier, 2009). The notion is that the TED spread may proxy for counterparty risk, which, when elevated, can lead to funding illiquidity. Our first hypothesis (H1) is that a positive shock to the *TED spread* is associated with a deterioration in systematic market efficiency.

As a more direct proxy for the availability of arbitrage capital, we include the total amount of assets under management at hedge funds in US$b. (*Hedge fund AUM*) in our analysis.[14] We expect greater assets under management at hedge funds to spur arbitrage activity, thus elevating the degree of systematic efficiency. Our second hypothesis (H2) is that a positive shock to *Hedge fund AUM* is associated with an improvement in systematic market efficiency.

[14] We thank Matti Suominen and LIPPER-TASS for data on hedge fund AUM. The sample includes all hedge funds that report their returns in U.S. dollars and have a minimum of 36 monthly return observations over our sample period. See Jylhä, Rinne, and Suominen (2015).

Furthermore, we include a proxy for algorithmic trading as a measure of the intensity of arbitrage and market making activity. Hendershott, Jones, and Menkveld (2011) and Brogaard, Hendershott, and Riordan (2014) provide evidence that algorithmic trading is associated with greater liquidity and more efficient pricing, which suggests that algorithmic or high-frequency traders play an important role in arbitrage and market making in today's markets. Consequently, more algorithmic trading should facilitate arbitrage, improving systematic efficiency. Inspired by Boehmer, Fong, and Wu (2015), our proxy for the intensity of algorithmic trading is defined as the total number of quote updates per month across all the stocks in our sample divided by the aggregate dollar trading volume for those stocks in the same month (*Quotes/Volume*). Our third hypothesis (H3) is that a positive shock to *Quotes/Volume* is associated with an improvement in systematic market efficiency.

Since market volatility is another potentially important determinant of the efficacy of market making and arbitrage, we include *Volatility* (computed as the value-weighted average across all stocks' average daily standard deviation of intraday one-minute mid-quote returns within the month) as another variable in the VAR, and our fourth hypothesis (H4) is that a positive shock to *Volatility* is associated with a deterioration in systematic market efficiency.

We further hypothesize that shocks to funding liquidity and the intensity of arbitrage activity may have a differential impact on the market efficiency of different sub-groups of stocks. In particular, Griffin and Xu (2009) show that hedge funds exhibit a strong preference for more liquid stocks, as measured by their turnover. Hence, our fifth hypothesis (H5) is that a positive shock to *Hedge fund AUM* is associated with a relatively greater improvement in systematic market efficiency for high turnover stocks.

4.2 Measuring systematic market efficiency

To test our hypotheses, we are interested in analyzing the relation of the *TED spread*, *Hedge fund AUM*, *Quotes/Volume*, and *Volatility* with the degree of systematic market efficiency. To this end, we extract a single, comprehensive measure of monthly, aggregate market efficiency via principal component analysis (PCA) of four of the five different monthly, market-wide efficiency measures: *Predictability*, *Hasbrouck*, *Put-call parity*, and *HKS*. We exclude *Variance ratio* from the principal component analysis because of its negative correlation with the other measures. However, our results are robust to including the *Variance ratio* measure. We follow Hasbrouck and Seppi (2000) and extract the principal components based on the correlation matrix. We find that the first principal component explains 39% of the total variation in the four market-wide efficiency measures.

Importantly, the loadings of the four different market-wide efficiency measures on the first principal component are all of the same sign, otherwise this component could not be interpreted as representing systematic variation in market efficiency. Since the loadings on the second principal component are not of the same sign, including this component in our systematic market efficiency measure would lead to problems in interpreting the resulting measure as being positively associated with the degree of efficiency as reflected in each of the four market-wide efficiency measures. Consequently, we use only the first principal component as representative of systematic market efficiency. The fact that this component explains almost 40% of the total variation lends credibility to the view that this component captures the dominant variation in systematic market efficiency. The loading of the first principal component on the underlying efficiency measures is 0.59 for *Predictability*, 0.75 for *Hasbrouck*, 0.27 for *Put-call parity*, and 0.11 for *HKS*. The first principal component is thus reasonably representative of all four efficiency measures and is not dominated by a single one of these measures.

To obtain a time-series of the first principal component, we standardize each of the four market-wide efficiency measures to have zero mean and unit standard deviation, and multiply the matrix of standardized measures by the vector of the loadings of each measure on the component. We refer to the resulting measure as *AEFF* (aggregate market efficiency) in the remainder of the paper.

Figure 2 presents a graph of the monthly time-variation in this comprehensive measure of market efficiency. We note that, just like the four underlying market-wide efficiency measures, *AEFF* is inversely related to the degree of informational efficiency. The figure shows that the degree of aggregate efficiency is considerably greater in some periods than in others and shares three pervasive features of several of the individual market-wide efficiency measures in Figure 1: the gradual improvement in efficiency over the period 2000-2008, and the sudden deteriorations in efficiency around the LTCM crisis and the collapse of Lehman Brothers.

To examine whether shocks to funding liquidity and the intensity of arbitrage activity have a differential impact on the aggregate market efficiency of low and high turnover stocks, we also estimate separate systematic efficiency components for two sub-groups of stocks sorted by turnover. At the beginning of each year, we sort stocks with below-median turnover over the previous year into a low turnover portfolio and stocks with above-median turnover into a high turnover portfolio. We then obtain a systematic market efficiency component for the low and high turnover portfolios separately in an analogous way as *AEFF* for all stocks. The first principal component explains close to 40% of the total variation in the different portfolio-level efficiency measures for both low and high turnover stocks.

4.3 Vector autoregression specification

A vector autoregression (VAR) is a natural way to analyze the dynamics of *AEFF* in relation to proxies for funding liquidity and the intensity of arbitrage activity as well as market volatility, since all of these variables are endogenous and could influence each other both contemporaneously and with a lag. We thus estimate a VAR with the following five endogenous variables: *TED spread*, *Hedge fund AUM*, *Quotes/Volume*, *Volatility*, and *AEFF*. We also estimate two additional VARs with the systematic efficiency components of low and high turnover stocks as key endogenous variables (as described in Section 4.2). For these other VARs, we re-estimate *Quotes/Volume* and *Volatility* based on the subsamples of low and high turnover stocks. Prior to usage as endogenous variables in the VAR, we detrend all five variables in each of the three VARs with linear and quadratic trend terms (to preclude spurious results).[15]

The number of lags in the VARs is determined using the Akaike and Schwarz information criteria (AIC and SIC). The AIC indicates 2, 12, and 10 lags for the VARs based on all stocks, low turnover stocks, and high turnover stocks, respectively, while the SIC indicates one lag for each of the three VARs. For the sake of parsimony, we choose to report the results of one-lag VARs (as indicated by the SIC), but VARs using the lag lengths indicated by the AIC yield similar results. Table 5 presents summary statistics of the four potential determinants of *AEFF*.

4.4 Empirical results on economic drivers of systematic efficiency

Table 6 reports contemporaneous correlations between the innovations (residuals) in the five endogenous variables in each of the three VARs: based on all stocks (Panel A), low

[15] The Augmented Dickey Fuller test rejects the null-hypothesis of a unit root for all variables included in the three VARs with p-values below 0.01.

turnover stocks (Panel B), and high turnover stocks (Panel C). Panel A shows strong positive contemporaneous correlations between innovations in *AEFF* and innovations in the *TED spread* and *Volatility*, which suggests that reduced overall funding liquidity and heightened volatility are associated with a deterioration in aggregate efficiency across all stocks (consistent with hypotheses H1 and H4). In line with hypothesis H2, innovations in *AEFF* are negatively correlated with innovations in *Hedge fund AUM*. These correlations are also economically substantial. There is no significant contemporaneous correlation between innovations in *AEFF* and innovations in *Quotes/Volume*.

Panel A of Table 6 also shows that innovations in *Volatility* are positively and significantly correlated with innovations in *Quotes/Volume* and the *TED spread*, and negatively and significantly with innovations in *Hedge fund AUM*. Somewhat surprisingly, correlations of innovations in *Quotes/Volume* are significantly negative with innovations in *Hedge fund AUM* and significantly positive with innovations in the *TED spread*, respectively, suggesting that algorithmic trading activity might intensify in times when hedge fund assets under management are reduced and funding liquidity deteriorates. The table further shows a significantly negative contemporaneous correlation between innovations in the *TED spread* and innovations in *Hedge fund AUM*, which conforms to our intuition that a reduction in overall funding liquidity is associated with a decline in the amount of arbitrage capital.

Panels B and C of Table 6 show very similar patterns in the residual correlations of the five variables included in the VARs based on low and high turnover stocks, respectively. A minor difference is that the correlation between innovations in *AEFF* and innovations in the *TED spread* is somewhat weaker for high turnover stocks.

Although the residual correlations in Table 6 provide some initial evidence on the relations between the endogenous variables in the VARs, they do not account for the full dynamics of the VAR system, and for the fact that shocks to the different endogenous variables are correlated. Impulse response functions (IRFs) provide a more complete picture by tracing the impact of a one time, unit standard deviation, orthogonalized (using the inverse Cholesky decomposition) shock to one of the endogenous variables on current and future values of the other endogenous variables.

Computation of Cholesky IRFs involves ordering the variables. In the ordering, the notion is that variables are affected by contemporaneous shocks to other variables that are above, but not below, them in the ordered list (see, for example, Bruno and Shin, 2015). Since *AEFF* is our key variable of interest, we include it as the last variable in the VAR, so that it can be affected by contemporaneous shocks to all other variables. The ordering of the other four variables is determined in the following way, based on the general principle that slow-moving variables should be ordered before fast-moving variables. Since shocks to the *TED spread* may be expected to have a contemporaneous effect on the other two proxies for funding liquidity and arbitrage activity as well as on market volatility and aggregate market efficiency, it is the first variable in our ordering. We consider *Hedge fund AUM* to be a more direct proxy for variation in the amount of capital that is available to arbitrageurs, and thus include it as the next variable in our VAR, so that it can be influenced contemporaneously by shocks to the *TED spread*. As *Quotes/Volume* is arguably a proxy for actual arbitrage activity that could be influenced contemporaneously by variation in the availability of arbitrage capital as picked up by the *TED spread* and *Hedge fund AUM*, we include it after these two variables. Since market volatility may be directly influenced by the funding liquidity proxies and may in turn affect aggregate efficiency, we include *Volatility* as the fourth variable in the VAR (just before *AEFF*).[16]

[16]We obtain similar results when we include either aggregate market-wide trading volume or market returns (in excess of the risk-free rate) as additional endogenous variables in the VARs (ordered in front of *Volatility*).

Figure 3 presents IRFs for the VAR estimated with the systematic efficiency component across all stocks as key endogenous variable. The IRFs show the response (measured in standard deviations) of the variable mentioned in the vertical legend to the right of the figure to a Cholesky one standard deviation shock to the variable mentioned in the horizontal legend at the top of the figure. Each IRF graph shows the cumulative response up to three months ahead (solid line labeled "coef"; month 0 on the horizontal axis of the IRF graphs corresponds to the contemporaneous response), as well as the bootstrapped 90% confidence bands based on 1,000 runs (dashed lines labeled "lower" and "upper").

The main result in Figure 3 is that shocks to all three proxies for funding liquidity and the intensity of arbitrage activity have a significant effect on *AEFF*. First, consistent with hypothesis H1, the cumulative response of *AEFF* to a shock to the *TED spread* (top right IRF in Figure 3) is positive and significant both contemporaneously and at all three lags depicted in the graph and is economically meaningful, at 0.2 to 0.4 standard deviations. Second, the cumulative response of *AEFF* to a shock to *Hedge fund AUM* (fourth IRF on top row) is significantly negative, both contemporaneously and with a one-month lag. This effect is in line with hypothesis H2 and, at around 0.2 standard deviations, non-trivial in magnitude. Third, the cumulative response of *AEFF* to a shock to *Quotes/Volume* (third IRF on top row) is significantly negative at all horizons depicted in the graph and economically large (at more than 0.3 standard deviations), supportive of hypothesis H3. These results provide broad support for the prediction of recent theories (e.g., Gromb and Vayanos, 2002; Brunnermeier and Pedersen, 2009; Gârleanu and Pedersen, 2011; Liu and Mello, 2011) that an improvement in funding liquidity enhances the overall degree of market efficiency.

In line with H4, we also find a significantly positive response of *AEFF* to a shock to *Volatility* (second IRF in the top row). In the IRFs that depict the response of *Volatility* to shocks in the other endogenous variables in the VAR (second row of Figure 3), we observe a significantly positive and large response to a shock to either *Quotes/Volume* or to the *TED spread*, and a significantly negative response to a shock to *Hedge fund AUM*, suggesting that algorithmic trading and funding illiquidity exacerbate volatility, while hedge fund activity tends to moderate volatility. In the third and fourth rows of Figure 3, we find that a shock to the *TED spread* has a significantly positive effect on *Quotes/Volume* (consistent with Table 6), but a significantly negative effect on *Hedge fund AUM*. The latter result indicates that funding illiquidity reduces the amount of arbitrage capital available to hedge funds.

Table 7 tabulates the cumulative impulse responses in the VAR for all stocks, and compares them to the impulse responses of the VARs estimated separately based on low and high turnover stocks. To save space, the table only reports the responses of *AEFF* to shocks to the other four endogenous variables. Consistent with hypothesis H5, the table shows that the negative response of *AEFF* to a shock to *Hedge fund AUM* is considerably stronger (both in economic and statistical terms) for high turnover stocks than for low turnover stocks. The effect of *Hedge fund AUM* on the aggregate efficiency of low turnover stocks is also negative but not statistically significant at any of the horizons considered. As a more direct test of hypothesis H5 (see the online appendix), we also estimate a separate VAR with the *difference* in the aggregate efficiency between high and low turnover stocks as the efficiency-related endogenous variable, and find that a positive shock to *Hedge fund AUM* is associated with a significant improvement in the aggregate efficiency of high turnover stocks relative to low turnover stocks.

Table 7 also shows that the positive response of *AEFF* to a *Volatility* shock is roughly equally large and long-lasting for low turnover stocks, high turnover stocks, and all stocks. The negative effect of a shock to *Quotes/Volume* on *AEFF* in Figure 3 is only observed for high turnover stocks, and its economic and statistical significance are slightly larger than that observed for all stocks. We are not aware of studies that provide direct evidence on the preference of algorithmic traders for certain equity segments, but Hendershott, Jones, and Menkveld (2011) argue that their results on the effect of algorithmic trading on the liquidity of NYSE stocks over the period 2001-2005 are "consistent with the conventional wisdom that AT was more prevalent at the time for active, liquid stocks." (their p.18). On the other hand, the cumulative response of *AEFF* to a shock to the *TED spread* is stronger for low turnover stocks than for high turnover stocks, both from an economic and a statistical perspective, which suggests that funding liquidity is more crucial in enhancing efficiency within low turnover stocks.[17]

In sum, our VAR results are consistent with our central hypothesis that changes in funding liquidity and the intensity of arbitrage activity affect systematic variation in market efficiency. They also provide new empirical support for theoretical models that propose an important role for funding constraints as a determinant of the efficacy of arbitrage, and thus of the degree of market efficiency. Furthermore, they suggest that hedge funds and algorithmic traders tend to enhance market efficiency, especially for high turnover stocks.

5. Conclusions

Market efficiency remains central to the study of financial markets, but while several measures of stock price efficiency have been proposed, little is known about how these

[17]We also estimate four separate VARs with each of the individual efficiency measures underlying *AEFF* as key endogenous variable. The online appendix shows that the impulse responses exhibit largely similar patterns across measures, although we also observe some differences. For example, the negative effect of *Quotes/Volume* is most pronounced for the *Predictability* measure, while the negative effect of *Hedge fund AUM* is strongest for *Hasbrouck* and *HKS*. These findings could be indicative of different types of arbitrageurs focusing on different types of deviations from price efficiency. The positive effect of shocks to the *TED spread* obtains for three of the four different market-wide efficiency measures, but, somewhat surprisingly, reverses sign for *Predictability*. Also, the statistical significance of some of the individual responses for the different efficiency measures is attenuated relative to the effects observed for the overall *AEFF* measure in Figure 3 and Table 7, possibly as a result of the somewhat greater degree of noise in each of the individual measures.

measures vary over time, whether they co-move across stocks and with each other, and what economic forces drive time-variation in systematic market efficiency. We address this gap in the literature by considering how different high-frequency market efficiency measures co-move across individual stocks and with each other, and by analyzing the determinants of the systematic market efficiency component across stocks and measures.

We show that four daily, stock-level market efficiency measures (intraday return predictability, variance ratios, Hasbrouck's (1993) pricing errors, and put-call parity deviations) demonstrate considerable cross-sectional and time-series variation and exhibit significant co-movement across stocks. Moreover, we find that correlations among four monthly, market-wide efficiency measures built up from these daily, stock-level measures as well as a fifth market-wide measure based on Heston, Korajczyk, and Sadka (2010) are mostly economically and statistically significant. These findings suggest that time-variation in these different market efficiency measures is characterized by a significant systematic component.

Motivated by recent theoretical and empirical research, we hypothesize that funding liquidity and the intensity of arbitrage activity are important economic forces driving the degree of systematic market efficiency. To test this hypothesis, we perform a vector autoregression which includes proxies for funding liquidity and arbitrage activity, market volatility, and the first principal component extracted from four monthly, market-wide efficiency measures. We show that shocks to funding liquidity (the TED spread) and to variables that more directly measure the intensity of arbitrage activity (hedge fund assets under management and a proxy for algorithmic trading) have a significant impact on the degree of systematic market efficiency. Overall, our results point to a significant, systematic, time-varying component in different price efficiency measures for individual stocks, and to a material role of funding liquidity and the intensity of arbitrage activity in driving fluctuations in this component.

Recognizing that market efficiency is dynamic and has a systematic component opens new vistas for research. First, it would be worth exploring whether there is *global* co-movement in market efficiency across stock markets in different countries. This would allow us to ascertain the extent to which the quality of price formation in markets across the world has a systematic component, and whether fluctuations in global funding liquidity affect the degree of global systematic efficiency. Second, it would be worth investigating whether systematic variation in market efficiency extends to other asset classes such as fixed income securities, foreign exchange, and derivatives. These and other issues are left for future research.

References

Amihud, Y., 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Andersen, T., T. Bollerslev, and A. Das, 2001, Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns, *Journal of Finance* 56, 305-327.

Andrade, S, C. Chang, and M. Seasholes, 2008, Trading imbalances, predictable reversals, and cross-stock price pressure, *Journal of Financial Economics* 88, 406-423.

Asquith, P., P. Pathak, and J. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243-276.

Barberis, N., A. Shleifer, and R. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.

Benston, G., and R. Hagerman, 1974, Determinants of bid-asked spreads in the over-the-counter market, *Journal of Financial Economics* 1, 353-364.

Berkman, H., P. Koch, L. Tuttle, and Y. Zhang, 2012, Paying attention: Overnight returns and the hidden cost of buying at the open, *Journal of Financial and Quantitative Analysis* 47, 715-741.

Bessembinder, H., 2003, Trade execution costs and market quality after decimalization, *Journal of Financial and Quantitative Analysis* 38, 747-777.

Beveridge, S., and C. Nelson, 1981, A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the measurement of the business cycle, *Journal of Monetary Economics* 7, 151-174.

Boehmer, E., K. Fong, and J. Wu, 2015, International evidence on algorithmic trading, working paper, Singapore Management University.

Boehmer, E., and E. Kelley, 2009, Institutional investors and the informational efficiency of prices, *Review of Financial Studies* 22, 3563-3594.

Boehmer, E., and J. Wu, 2007, Order flow and prices, working paper, University of Georgia.

Brogaard, J., T. Hendershott, and R. Riordan, 2014, High frequency trading and price discovery, *Review of Financial Studies* 27, 2267-2306.

Brunnermeier, M., 2009, Deciphering the liquidity and credit crunch 2007-2008, *Journal of Economic Perspectives* 23, 77-100.

Brunnermeier, M., S. Nagel, and L. Pedersen, 2008, Carry trades and currency crashes, *NBER Macroeconomics Annual*, Chapter 5, 313-347.

Brunnermeier, M., and L. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

Bruno, V., and H. Shin, 2015, Capital flows and the risk-taking channel of monetary policy, *Journal of Monetary Economics* 71, 119-132.

Chan, K., and W. Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, *Journal of Financial Economics* 57, 247-273.

Charles, A., and O. Darné, 2009, Variance-ratio tests of random walk: An overview, *Journal of Economic Surveys* 23, 503-527.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3-28.

Chordia, T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111-130.

Chordia, T., R. Roll, and A. Subrahmanyam, 2005, Evidence on the speed of convergence to market efficiency, *Journal of Financial Economics* 76, 271-292.

Chordia, T., R. Roll, and A. Subrahmanyam, 2008, Liquidity and market efficiency, *Journal of Financial Economics* 87, 249-268.

Cremers, M., and D. Weinbaum, 2010, Deviations from put-call parity and stock return predictability, *Journal of Financial and Quantitative Analysis* 45, 335-367.

D'Avolio, G, 2002, The market for borrowing stock, *Journal of Financial Economics* 66, 271-306.

Desai, H., K. Ramesh, S. Thiagarajan, and B. Balachandran, 2002, An investigation of the informational role of short interest in the Nasdaq market, *Journal of Finance* 57, 2263-2287.

Domowitz, I., J. Glen, and A. Madhavan, 1998, International cross-listing and order flow migration: Evidence from an emerging market, *Journal of Finance* 53, 2001-2027.

Fama, E., 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383-417.

Finucane, T., 1991, Put-call parity and expected returns, *Journal of Financial and Quantitative Analysis* 26, 445-457.

Foucault, T., R. Kozhan, and W. Tham, 2015, Toxic arbitrage, working paper, HEC Paris.

Gao, X., and J. Ritter, 2010, The marketing of seasoned equity offerings, *Journal of Financial Economics* 97, 33-52.

Gârleanu, N., and L. Pedersen, 2011, Margin-based asset pricing and deviations from the Law of One Price, *Review of Financial Studies* 24, 1980-2022.

Glosten, L., and P. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.

Goyenko, R., C. Holden, and C. Trzcinka, 2009, Do liquidity measures measure liquidity?, *Journal of Financial Economics* 92, 153-181.

Griffin, J., and J. Xu, 2009, How smart are the smart guys? A unique view from hedge fund stock holdings, *Review of Financial Studies* 22, 2531-2570.

Gromb, D., and D. Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361-407.

Gromb, D., and D. Vayanos, 2010, Limits of arbitrage, *Annual Review of Financial Economics* 2, 251-275.

Hameed, A., W. Kang, and S. Viswanathan, 2010, Stock market declines and liquidity, *Journal of Finance* 65, 257-293.

Hasbrouck, J., 1993, Assessing the quality of a security market: A new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191-212.

Hasbrouck, J., 2009, Trading costs and returns for U.S. equities: Estimating effective costs from daily data, *Journal of Finance* 64, 1445-1477.

Hasbrouck, J., and T. Ho, 1987, Order arrival, quote behavior, and the return-generating process, *Journal of Finance* 42, 1540-6261.

Hasbrouck, J., and D. Seppi, 2000, Common factors in prices, order flows and liquidity, *Journal of Financial Economics* 59, 383-411.

Hendershott, T., C. Jones, and A. Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1-33.

Heston, S., R. Korajczyk, and R. Sadka, 2010, Intraday patterns in the cross-section of stock returns, *Journal of Finance* 65, 1369-1407.

Holden, C., 1995, Index arbitrage as cross-sectional market making, *Journal of Futures Markets* 15, 423-455.

Huberman, G., and D. Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161-178.

Jones, C., and O. Lamont, 2002, Short sales constraints and stock returns, *Journal of Financial Economics* 66, 207?239.

Jylhä, P., K. Rinne, and M. Suominen, 2015, Do hedge funds supply or demand immediacy?, *Review of Finance*, forthcoming.

Karolyi, G.A., K.-H. Lee, and M. van Dijk, 2012, Understanding co-movement in liquidity around the world, *Journal of Financial Economics* 105, 82-112.

Kumar, P., and D. Seppi, 1994, Information and index arbitrage, *Journal of Business* 67, 481-509.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Lee, C., and B. Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, *Journal of Financial Markets* 3, 83-111.

Lee, C., and M. Ready, 1991, Inferring trade direction from intra-day data, *Journal of Finance* 46, 733-747.

Liu, X., and A. Mello, 2011, The fragile capital structure of hedge funds and the limits to arbitrage, *Journal of Financial Economics* 102, 491-506.

Lo, A., and C. MacKinlay, 1989, The size and power of the variance ratio test in finite samples: A Monte Carlo investigation, *Journal of Econometrics* 40, 203-238.

Mancini-Griffoli, T., and A. Ranaldo, 2011, Limits to arbitrage during the crisis: Funding liquidity constraints and covered interest parity, working paper, Swiss National Bank.

Mitchell, M., L. Pedersen, and T. Pulvino, 2007, Slow moving capital, *American Economic Review Papers and Proceedings* 97, 215-220.

Mitchell, M., and T. Pulvino, 2012, Arbitrage crashes and the speed of capital, *Journal of Financial Economics* 104, 469-490.

Mitchell, M., T. Pulvino, and E. Stafford, 2002, Limited arbitrage in equity markets, *Journal of Finance* 57, 551-584.

Morck, R., B. Yeung, and W. Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous stock price movements?, *Journal of Financial Economics* 58, 215-260.

Nagel, S., 2005, Short-sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277-309.

Nagel, S., 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005-2039.

Newey, W., and K. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* 61, 631-653.

Odders-White, E., 2000, On the occurrence and consequences of inaccurate trade classification, *Journal of Financial Markets* 3, 205-332.

Pan, J., 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3-50.

Pasquariello, P., 2014, Financial market dislocations, *Review of Financial Studies* 27, 1868-1914.

Shleifer, A., and R. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Stoll, H., 1978, The supply of dealer services in securities markets, *Journal of Finance* 33, 1133-1151.

Table 1 – Summary statistics of input variables for intraday return predictability regressions

This table reports the cross-sectional (across the 2,157 NYSE stocks in the sample) mean, standard deviation ("SD"), first quartile ("25%"), median, and third quartile ("75%") of the time-series average by stock of the daily number of trades (*#trades*), daily trading volume in US$ billions (*dollar volume*), average one-minute mid-quote returns within the day in basis points (*1-min mid-quote return*), average difference between the total number of trades initated by buyers and sellers (order imbalance in number of trades) over one-minute intervals (*1-min oib#*), and the average difference between the total dollar volume of trades initiated by buyers and sellers (order imbalance in US$) over one-minute intervals (*1-min oib$*). The first column indicates the number of stocks over which the summary statistics are computed. The sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data to compute all variables are from TRTH.

	#Stocks	Mean	SD	25%	Median	75%
#trades	2,157	2,015	4,798	79	432	1,797
dollar volume	2,157	0.025	0.062	0.001	0.006	0.021
1-min mid-quote return	2,157	-0.007	0.255	-0.023	-0.001	0.011
1-min oib#	2,157	0.067	0.138	0.001	0.023	0.094
1-min oib$	2,157	3,052	7,077	4	520	2,616

Table 2 – Intraday return predictability regressions and summary statistics of stock-level efficiency measures

Panel A of this table reports the average results of the return predictability regressions from Eq. (1), estimated daily based on intraday data for each of the NYSE stocks in the sample. Each of the five columns presents the results of a different way to estimate the predictability of one-minute (or two-minute) returns from lagged order imbalance (*OIB*) or lagged returns: *OIB predictability*, *allquotes*, *2minutes*, *oib#*, and *autocorrelation*. Section 1.1 discusses all five return predictability measures in detail. The first number in each column is the average slope coefficient across all stock-day predictability regressions. The *OIB* coefficient is scaled by 10^9 for the *OIB predictability*, *allquotes*, and *2minutes* regressions and by 10^4 for the *oib#* regressions. The average t-statistics ("t-stat avg") and the average Newey-West (1994) t-statistics ("NW t-stat avg") are reported below the coefficients. "% positive" is the percentage of positive coefficients, and "% + significant" is the percentage with t-statistics greater than 1.645 (the 5% critical level in a one-tailed test). Intercepts have been suppressed to conserve space. The last three rows report the average R^2 and adjusted R^2 across all regressions and the number of stock-day predictability regressions.

Panel B reports the cross-sectional mean, standard deviation ("SD"), first quartile ("25%"), median, and third quartile ("75%") of the time-series average by stock of the five different return predictability measures from Panel A as well as the other three daily, stock-level efficiency measures (*Variance ratio*, *Hasbrouck*, and *Put-call parity*). *Variance ratio* is the daily, absolute difference between one and 30 times the ratio of the variance estimated from one-minute mid-quote returns to the variance estimated from 30-minute mid-quote returns. *Hasbrouck* is the daily maximum of the absolute intraday pricing errors extracted from a decomposition of observed prices into efficient prices and a stationary component (Hasbrouck, 1993). *Put-call parity* is the end-of-day absolute difference between the implied volatilities of near-the-money call and put options of the same series (i.e., pairs of options on the same underlying stock with the same strike price and the same expiration date). Section 1 discusses all stock-level efficiency measures in detail. The first column indicates the number of stocks over which the summary statistics are computed.

Panel C reports average cross-sectional Spearman rank correlations (in %) across the monthly, stock-level *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity* measures. *Predictability* is the common factor extracted via principal component analysis by stock of the five intraday return predictability measures from Panel A. The monthly, stock-level efficiency measures are constructed from the corresponding daily, stock-level measures by averaging across days within the month. Panel C reports p-values, based on Newey-West (1994) standard errors with automatic lag selection, in parentheses below the correlations. Significance at the 1%, 5%, and 10% level is indicated by ***, **, and *, respectively.

The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data are from TRTH. Data to compute *Put-call parity* are from OptionMetrics.

Panel A: Intraday return predictability regressions (with dependent variable $Return_t$)

Predictability measure:	*OIB predictability*	*allquotes*	*2minutes*	*oib#*	*autocorrelation*
OIB_{t-1}	4.380	3.792	6.169	0.947	
$Return_{t-1}$					0.024
t-stat avg	1.254	0.779	0.863	1.852	0.370
NW t-stat avg	8.385	7.042	7.800	10.657	3.614
% positive	81.61	72.03	74.62	88.71	58.00
% + significant	45.75	31.85	35.10	61.67	28.82
R^2 (in %)	2.55	1.72	2.57	3.47	1.83
adj. R^2 (in %)	1.60	0.75	1.45	2.51	0.79
# regressions	3,175,645	3,263,908	3,391,138	3,271,527	3,097,085

Table 2 – continued

Panel B: Summary statistics of daily, stock-level efficiency measures

	#Stocks	Mean	SD	25%	Median	75%
OIB predictability	1,924	13.324	34.030	1.183	3.258	9.518
allquotes	1,924	8.714	24.971	0.598	1.538	5.263
2minutes	1,956	12.602	29.267	1.062	3.140	9.466
oib#	1,919	1.652	2.090	0.675	0.973	1.710
autocorrelation	1,837	0.023	0.033	0.003	0.023	0.044
Variance ratio	2,130	0.867	0.378	0.632	0.762	0.971
Hasbrouck	1,769	0.394	0.443	0.153	0.245	0.458
Put-call parity	1,535	2.581	2.043	1.467	2.091	3.073

Panel C: Cross-sectional correlations across monthly, stock-level efficiency measures

	Predictability	*Variance ratio*	*Hasbrouck*	*Put-call parity*
Predictability	100.00			
Variance ratio	3.22 (0.26)	100.00		
Hasbrouck	35.95*** (0.00)	21.91*** (0.00)	100.00	
Put-call parity	25.63*** (0.00)	22.43*** (0.00)	52.69*** (0.00)	100.00

Table 3 – Co-movement regressions of daily, stock-level efficiency on market efficiency

This table reports the average results of the efficiency co-movement regressions from Eq. (2), estimated for each NYSE stock in the sample based on daily data over the whole sample period. The dependent variable $Eff_{i,d}$ is the efficiency of stock i on day d, orthogonalized with respect to stock i's proportional quoted spread (*PQSPR*) on day d. The independent variable $MktEff_d$ is the (orthogonalized) market-wide efficiency on day d, computed as the value-weighted average efficiency (orthogonalized with respect to *PQSPR*) of all individual stocks on day d, excluding stock i. Each co-movement regression also includes a one-day lead and lag of (orthogonalized) market-wide efficiency. Each of the four columns in the table presents the results of the co-movement regressions based on a different stock-level efficiency measure: *Predictability*, *Variance ratio*, *Hasbrouck*, and *Put-call parity*. *Predictability* is the common factor extracted via principal component analysis by stock of the five intraday return predictability measures from Panel A of Table 2. We refer to Table 2 for a description of these predictability measures as well as the other three stock-level efficiency measures. Each column presents the average slope coefficients across all co-movement regressions estimated by stock. The average t-statistics ("t-stat avg") and the cross-sectional t-statistics ("CS t-stat") are reported below the coefficients. "% positive" is the percentage of positive coefficients, and "% + significant" is the percentage with t-statistics greater than 1.645 (the 5% critical level in a one-tailed test). Intercepts have been suppressed to conserve space. The last three rows report the average R^2 and adjusted R^2 across all regressions, and the number of stocks for which the co-movement regressions were estimated. The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data are from TRTH. Data to compute *Put-call parity* are from OptionMetrics.

	Dependent variable: $Eff_{i,d}$			
Efficiency measure:	*Predictability*	*Variance ratio*	*Hasbrouck*	*Put-call parity*
$MktEff_d$	0.483	0.699	1.528	1.064
t-stat avg	5.981	5.482	4.646	4.835
CS t-stat	5.151	4.665	2.068	2.880
% positive	87.58	92.21	88.54	90.23
% + significant	69.30	72.46	66.00	68.44
$MktEff_{d-1}$	0.033	0.080	0.147	0.176
t-stat avg	0.534	0.123	0.309	0.844
CS t-stat	0.428	0.825	0.361	0.631
% positive	57.35	51.40	56.86	61.20
% + significant	27.71	11.66	18.53	28.37
$MktEff_{d+1}$	0.033	0.098	-0.061	0.180
t-stat avg	0.712	0.064	0.313	0.834
CS t-stat	0.507	1.019	-0.106	0.560
% positive	58.76	50.81	59.30	62.06
% + significant	30.70	10.44	19.10	28.64
R^2 (in %)	6.47	3.23	6.73	8.81
adj R^2 (in %)	5.21	2.31	4.94	7.96
# regressions	1,707	2,041	1,597	1,505

Table 4 – Time-series correlations across monthly, market-wide efficiency measures

This table reports Spearman rank correlations (in %) across five monthly, market-wide efficiency measures: *Predictability*, *Variance ratio*, *Hasbrouck*, *Put-call parity*, and *HKS*. The first four of these measures are aggregated from daily, stock-level efficiency measures by averaging the daily, stock-level efficiency measures across days within the month to construct monthly, stock-level efficiency measures, and then computing the value-weighted average efficiency across individual stocks each month, separately for each efficiency measure. The HKS measure is obtained from regressing 30-minute stock returns on returns over the same interval on the previous day (averaged over all 30-minute intervals within a month). The table reports p-values in parentheses below the correlations. The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data are from TRTH. Data to compute *Put-call parity* are from OptionMetrics. Significance at the 1%, 5%, and 10% level is indicated by ***, **, and *, respectively.

	Predictability	*Variance ratio*	*Hasbrouck*	*Put-call parity*	*HKS*
Predictability	100.00				
Variance ratio	-36.64***	100.00			
	(0.00)				
Hasbrouck	59.75***	-35.78***	100.00		
	(0.00)	(0.00)			
Put-call parity	18.83**	-37.10***	42.07***	100.00	
	(0.01)	(0.00)	(0.00)		
HKS	-3.49	-3.58	14.87**	-3.66	100.00
	(0.64)	(0.64)	(0.05)	(0.63)	

Table 5 – Summary statistics of potential determinants of systematic variation in efficiency

This table reports the number of monthly observations and the time-series mean, standard deviation ("SD"), first quartile ("25%"), median, and third quartile ("75%") of four potential determinants of monthly variation in the systematic component of market efficiency. *TED spread* is the monthly difference between the three-month LIBOR and the three-month T-bill rate (in %), obtained from the FRED database of the Federal Reserve Bank of St. Louis (FRED ID: USD3MTD156N minus TB3MS). *Hedge fund AUM* is the total amount of assets under management at hedge funds in US$b., obtained from Matti Suominen and LIPPER-TASS (see Jylhä, Rinne, and Suominen, 2015). *Quotes/Volume* is the total number of quote updates per month across all the NYSE stocks in our sample divided by the aggregate dollar trading volume for those stocks in the same month. This variable is scaled by 10^2. *Volatility* is a measure of the overall volatility of the stock market, computed as the value-weighted average across all stocks' average daily standard deviation of intraday one-minute mid-quote returns within the month. Data to compute *Quotes/Volume* and *Volatility* are from TRTH. The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix).

	# Obs.	Mean	SD	25%	Median	75%
TED spread	177	0.576	0.443	0.241	0.484	0.729
Hedge fund AUM	177	905.76	525.86	443.39	740.74	1414.80
Quotes/Volume	177	0.02	0.02	0.00	0.01	0.03
Volatility	177	31.47	10.76	24.75	28.61	35.98

Table 6 – Vector autoregressions of systematic market efficiency (*AEFF*): Residual correlations

This table reports the contemporaneous correlations (in %) between the innovations (residuals) in the following five endogenous variables included in three separate vector autoregressions (VARs) based on all NYSE stocks in the sample (Panel A), based on low turnover stocks only (Panel B), and based on high turnover stocks only (Panel C): the TED spread (*TED spread*), hedge funds assets under management (*Hedge fund AUM*), the total number of quote updates scaled by aggregate dollar trading volume (*Quotes/Volume*), market volatility (*Volatility*), and the systematic market efficiency component or "aggregate market efficiency" (*AEFF*). We refer to Table 5 and Figure 2 for a description of these variables. The aggregate market efficiency measures (*AEFF*) based on low and high turnover stocks separately are constructed in an analogous way as *AEFF* for all stocks. For the VARs in Panels B and C, we also re-estimate *Quotes/Volume* and *Volatility* based on low and high turnover stocks only. The classification of all NYSE stocks into low and high turnover stocks is based on the median turnover over the previous year. We estimate all three VARs with one lag, following the Schwarz information criterion (SIC). All variables in the VARs have been detrended using a linear and a quadratic time trend. The table reports p-values in parentheses below the correlations. Significance at the 1%, 5%, and 10% level is indicated by ***, **, and *, respectively.

Panel A: All stocks

	AEFF	*Volatility*	$\frac{Quotes}{Volume}$	*Hedge fund AUM*	*TED spread*
AEFF	100.00				
Volatility	44.85***	100.00			
	(0.00)				
Quotes/Volume	-2.51	59.07***	100.00		
	(0.74)	(0.00)			
Hedge fund AUM	-25.59***	-42.89***	-22.35***	100.00	
	(0.00)	(0.00)	(0.00)		
TED spread	23.02***	52.71***	41.09***	-27.82***	100.00
	(0.00)	(0.00)	(0.00)	(0.00)	

Table 6 – continued

Panel B: Low turnover stocks

	AEFF	*Volatility*	$\frac{Quotes}{Volume}$	*Hedge fund AUM*	*TED spread*
AEFF	100.00				
Volatility	52.01***	100.00			
	(0.00)				
Quotes/Volume	7.55	47.91***	100.00		
	(0.32)	(0.00)			
Hedge fund AUM	-20.78***	-40.27***	-16.50**	100.00	
	(0.01)	(0.00)	(0.03)		
TED spread	31.02***	52.08***	26.65***	-27.21***	100.00
	(0.00)	(0.00)	(0.00)	(0.00)	

Panel C: High turnover stocks

	AEFF	*Volatility*	$\frac{Quotes}{Volume}$	*Hedge fund AUM*	*TED spread*
AEFF	100.00				
Volatility	50.98***	100.00			
	(0.00)				
Quotes/Volume	-6.78	46.38***	100.00		
	(0.37)	(0.00)			
Hedge fund AUM	-34.92***	-45.26***	-20.63***	100.00	
	(0.00)	(0.00)	(0.01)		
TED spread	14.49*	50.95***	46.28***	-28.55***	100.00
	(0.05)	(0.00)	(0.00)	(0.00)	

Table 7 – VARs of systematic market efficiency (*AEFF*): Impulse response functions

This table reports the impulse response functions (IRFs) for three separate vector autoregressions (VARs) with one lag based on all NYSE stocks in the sample (Panel A), based on low turnover stocks only (Panel B), and based on high turnover stocks only (Panel C), with the following endogenous variables (in this order): *TED spread*, *Hedge fund AUM*, *Quotes/Volume*, *Volatility*, and the systematic component of market efficiency *AEFF* (see variable descriptions in Tables 5-6 and Figure 2). The endogenous variables *AEFF*, *Quotes/Volume*, and *Volatility* are each constructed separately for each Panel based on the relevant sample of stocks. Each Panel shows the cumulative response of *AEFF* to a Cholesky one standard deviation shock to the variable in each row. To save space, the table does not report the impulse responses of the other four endogenous variables in each VAR. Responses, shown from month 0 (contemporaneous) up to 3 months after the shock, are measured in standard deviations. The impulse responses in Panel A are the same as those in the top row of IRFs of Figure 3. Significance at the 5% and 10% level is indicated by ** and *, respectively.

AEFF across:	**Panel A: All stocks**	**Panel B: Low turnover stocks**	**Panel C: High turnover stocks**
$AEFF_m$	0.610**	0.685**	0.546**
$AEFF_{m-1}$	0.979**	1.009**	0.929**
$AEFF_{m-2}$	1.204**	1.171**	1.204**
$AEFF_{m-3}$	1.344**	1.259**	1.402**
$Volatility_m$	0.349**	0.369**	0.344**
$Volatility_{m-1}$	0.616**	0.534**	0.589**
$Volatility_{m-2}$	0.809**	0.623**	0.759**
$Volatility_{m-3}$	0.943**	0.683**	0.874**
$Quotes/Volume_m$	-0.118**	-0.017	-0.127**
$Quotes/Volume_{m-1}$	-0.211**	0.017	-0.268**
$Quotes/Volume_{m-2}$	-0.289*	0.065	-0.402**
$Quotes/Volume_{m-3}$	-0.355*	0.110	-0.519**
$Hedge\ fund\ AUM_m$	-0.150*	-0.106	-0.226**
$Hedge\ fund\ AUM_{m-1}$	-0.235*	-0.148	-0.380**
$Hedge\ fund\ AUM_{m-2}$	-0.278	-0.161	-0.485**
$Hedge\ fund\ AUM_{m-3}$	-0.291	-0.157	-0.554**
$Ted\ spread_m$	0.172**	0.256**	0.102*
$Ted\ spread_{m-1}$	0.283**	0.450**	0.186*
$Ted\ spread_{m-2}$	0.358*	0.600**	0.247*
$Ted\ spread_{m-3}$	0.406*	0.642**	0.285

Figure 1 – Monthly variation in individual market-wide efficiency measures

This figure shows monthly variation in five different market-wide efficiency measures: *Predictability*, *Variance ratio*, *Hasbrouck*, *Put-call parity*, and *HKS*. We refer to Table 4 for a description of these measures. Each measure is an inverse indicator of the degree of market efficiency. The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data are from TRTH. Data to compute *Put-call parity* are from OptionMetrics.

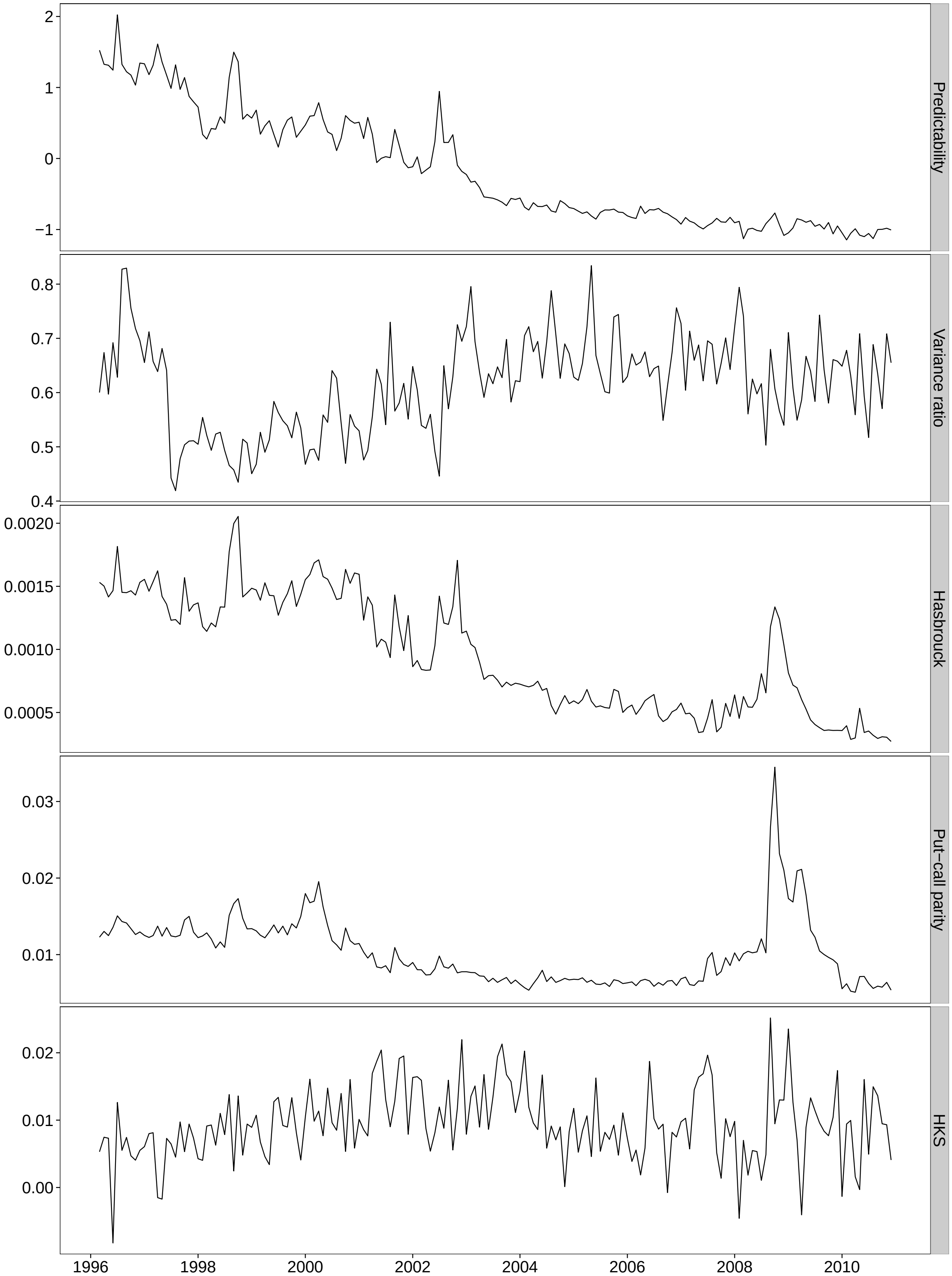

Figure 2 – Monthly variation in systematic market efficiency (*AEFF*)

This figure shows monthly variation in the systematic component of market efficiency across stocks and measures, or "aggregate market efficiency" (*AEFF*). *AEFF* is the common factor extracted via principal component analysis based on four of the five monthly, market-wide efficiency measures in Figure 1: *Predictability*, *Hasbrouck*, *Put-call parity*, and *HKS*. We refer to Table 4 for a description of these measures. *AEFF* is an inverse indicator of the degree of aggregate market efficiency. The full sample includes all 2,157 NYSE-listed common stocks from 1996 to 2010 that survive our data screens (described in the online appendix). Data are from TRTH. Data to compute *Put-call parity* are from OptionMetrics.

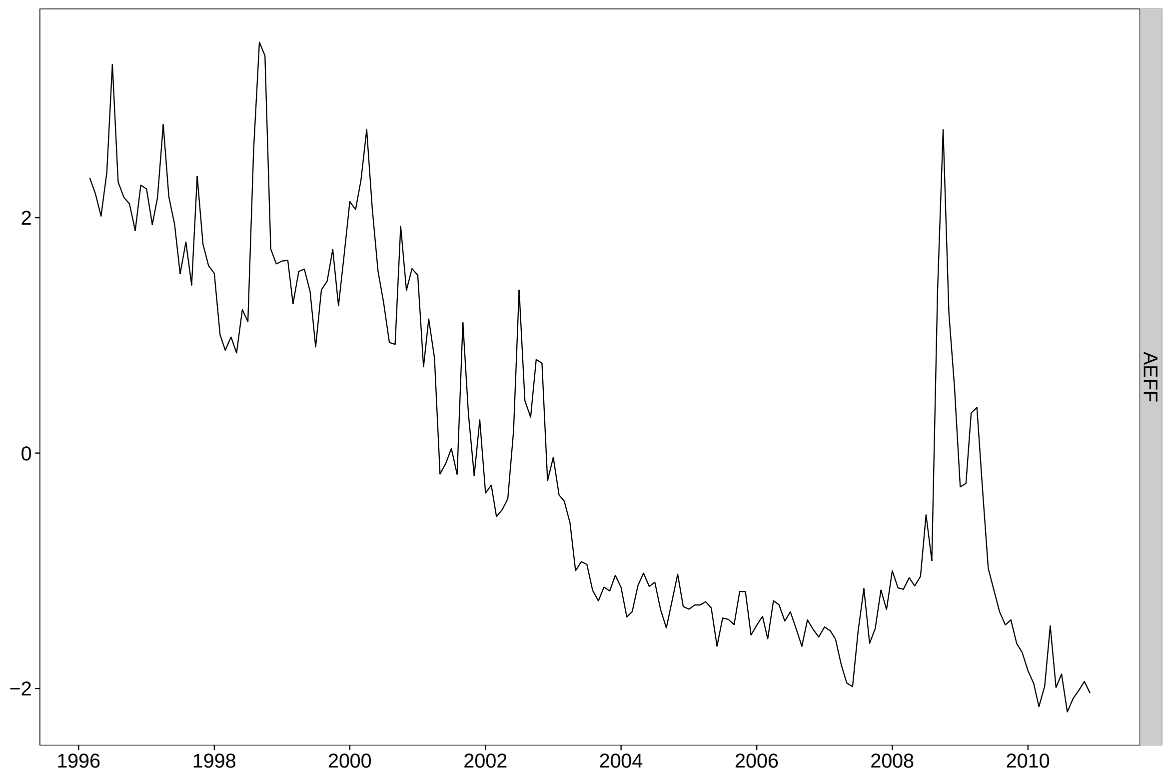

Figure 3 – VAR of systematic market efficiency (*AEFF*): Impulse response functions

This figure shows impulse response functions (IRFs) for a vector autoregression (VAR) with one lag based on all NYSE stocks, with the following endogenous variables (in this order): *TED spread*, *Hedge fund AUM*, *Quotes/Volume*, *Volatility*, and the systematic component of market efficiency *AEFF* (see variable descriptions in Table 5 and Figure 2). Each IRF shows the cumulative response ("coef") of the variable in the vertical legend to the right of the figure to a Cholesky one standard deviation shock to the variable in the horizontal legend at the top of the figure, and bootstrapped 90% confidence bands ("lower" and "upper"). Responses (measured in standard deviations) are shown from month 0 (contemporaneous) up to 3 months after the shock.