We study the dynamics of high-frequency market efficiency measures. We provide evidence that these measures comove 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. (JEL G10, G14)

Received April 30, 2016; editorial decision June 27, 2016 by Editor Stefan Nagel

We thank Stefan Nagel (the editor), two anonymous referees, Yakov Amihud, Tarun Chordia, Carole Comerton-Forde, Thierry Foucault, Amit Goyal, Terry Hendershott, Craig Holden, Seemti Kamma, Andrew Karolyi, Ed Lin, Marc Lipson, Steve Mann, Christophe Pérignon, Veronika Pool, Vikas Raman, Raghun 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 Bensheim School of Management), the 2013 EFMA 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 SURFdsara, and in particular Lykle Veert, 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. Supplementary data can be found on The Review of Financial Studies web site. Send correspondence to Dominik Rösch, State University of New York at Buffalo, 366 Jacobs Management Center, NY 14260; telephone: +1 716 645 9095. E-mail: drosch@buffalo.edu.

© The Author 2016. Published by Oxford University Press on behalf of The Society for Financial Studies. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

markets, but it does not allow for market efficiency itself to vary through time, in predictable and in 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 price convergence on 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) for which severity considerably varies 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 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 largely have been separately investigated in the literature. However, we note that they are all intimately linked to arbitrage and market making, both of which are impeded by time-varying financial frictions. And although these frictions differ across individual securities, they also have a systematic component. Thus, there may be a significant systematic component to the time-varying behavior of market efficiency measures.

Motivated by the above observations, we ask the following questions. To what extent do different market efficiency measures vary over time? Do different market efficiency measures comove 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 systematic variation. For example, investors’ allocations to equities may be influenced by their systematic degree of price efficiency. Moreover, researchers could benefit

1 See, for example, Shleifer and Vishny (1997), Mitchell, Palvino, and Setteferr (2009), and Aveta (2008) for theoretical and empirical explorations of how limits to arbitrage can cause market inefficiencies to persist.

2 See, for example, Brennan and Hamermesh (1989) and Wang (2008) for evidence on cross-sectional variation in stock-level illiquidity and short-sales constraints, respectively. See, for example, Hasbrouck and Seppi (2008) 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 2006), as well as macroeconomic funding constraints (Hammermeister and Pedersen 2009). Keane and Pedersen (2001) link deviations from the law of one price to variation in the aggregate shadow cost of capital. Desai et al. (2008), Jones and Lamont (2000), and Asquith, Pashak, and Ritter (2004) provide evidence of systematic variation in short-sale constraints.
The Dynamics of Market Efficiency

from a better understanding of the extent to which the different efficiency measures used in the literature are related and of whether these measures 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’s 2003), intraday Hasbrouck (1993) pricing errors, and put-call parity deviations in the corresponding options markets (Finucane 1991; Cremers and Weinbaum 2010). We compute these measures by 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.3

We construct market-wide measures of efficiency from each of the stock-level measures and estimate the degree of “comovement 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; this indicates that market efficiency measures are prone to systematic improvement and deterioration.

We then examine comovement 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 (2014). These correlations are mostly economically substantial and statistically significant, with the correlations of the variance ratio measure vis-à-vis the other efficiency measures being a notable exception. This finding indicates that four of the five market-wide efficiency measures share significant common variation, a fact that 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; Gromb and Vayanos 2002; Brunnermeier and Pedersen 2009; Garleanu 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

3 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.
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 the component as the main variable of interest in vector autoregressions (VARs). 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] as other endogenous variables. 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 comove 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 have an impact 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
The Dynamics of Market Efficiency

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 long-standing 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 comove 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 (1993) pricing errors, and put-call parity deviations. Further, in our analysis of comovement across aggregate efficiency measures, we include the market-wide cross-sectional return predictability measure of Heston, Korajczyk, and Sadka (2010).

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 those by Hasbrouck and Hl (1987), Chan and Fong (2001), 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 in 1996 and 2002, respectively, cease to be predictable from order flow in thirty minutes or less and in around five minutes. Since our sample period lasts until 2010, it is judicious to use intervals shorter than five minutes to 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:

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

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 Equation (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 on 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 on past order flows, and label this the autocorrelation measure. We discard stock-days with fewer than twenty observations for each of these measures. In our analyses of comovement in market efficiency, we use a comprehensive Predictability measure 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’s (2003) measure. 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
The Dynamics of Market Efficiency

variance estimated from thirty-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 nonsynchronous trading. Since estimates of daily variance ratios of individual stocks can be noisy (Andersen, Bollerslev, and Das 2001), we follow Lo and MacKinlay (1988; see their equation (5)) and Charles and Darné (2005) 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 twenty nonzero 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 Hashbrouck pricing errors
As a third daily stock-level efficiency measure, we estimate Hashbrouck (1993) pricing errors based on intraday trades and quotes. Hashbrouck 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 and argues it is a natural measure for price inefficiency. We follow Hashbrouck and use 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: (1) the logarithmic price return, from quote midpoints associated with trades, (2) a trade sign indicator, (3) the signed volume (that is, the sign of the trade times the number of shares traded), and (4) 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 Hashbrouck 1993). As in Hashbrouck (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 given day from the vector moving average representation of the VAR system (Beveridge and Nelson 1981) using Equation (13) in Hashbrouck (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 Hashbrouck measure. Since daily stock-level

4 Using mid-quote returns avoids the bid-ask bounce, but using returns from actual trade prices does not alter the main results.
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 comovement 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). 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 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 comovement 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 thirty-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 not fully anticipated by the market. Following their approach, we divide the 6.5-hour trading day into thirteen thirty-minute intervals and run cross-sectional regressions of thirty-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 across the different market efficiency measures is how these measures are related to market liquidity. This measure is also used by Cremers and Weißenant (2010), who 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.
liquidity. Characterizing the relation between efficiency and liquidity is not straightforward, since the causality can run either way, competing hypotheses predict opposite-sign relations, and 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, alternative channels 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’s (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, 193–4). Hasbrouck (1993) does not explicitly consider market makers’ potential cognitive limitations, but they 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 market or the options market. Further, HK5 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 Sepp 1994), “poor intermarket information linkages” (Domowitz, Glen, and Madhavan 1998), or “stale quotes” (Foucault, Kozhan, and Than 2015). These three papers.
study inefficiencies in the context of, respectively, index arbitrage, cross-listings arbitrage, and triangular arbitrage in foreign exchange markets. The informational differences channel thus may be most directly relevant for law of one price deviations, and 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 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 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. They may apply to a lesser or a 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 considerably overlap with illiquidity, our analysis of comovement in efficiency across stocks might be
perceived as a reiteration of the extensive literature on comovement in liquidity (e.g., Chordia, Roll, and Subrahmanyam 2000; Hasbrouck and Seppä 2003; Huberman and Hakl 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.

Which 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 comovement 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 Sections 3.2 and 4, we orthogonalize the monthly stock-level \textit{Predictability}, \textit{Variance ratio}, \textit{Hasbrouck}, and \textit{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 \textit{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. 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, for example, 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, and these are described in the Online Appendix. Our final sample includes 2,157 NYSE stocks.

To estimate the predictability regressions in Equation (1), we require at least one signed trade in both the interval over which we calculate the return and over the previous interval. We discard stock-days for which we have fewer than twenty 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 datum filters allow us to estimate Equation (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.

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 25 million USD. The median one-minute mid-quote return is equal to $-0.001$ basis point, and this 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 et al. (2013) report negative mean and median open-to-close

---

6 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.

7 The Appendix also presents results from the robustness checks mentioned within the paper.

8 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 Ramakrishnan (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.
The Dynamics of Market Efficiency

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

<table>
<thead>
<tr>
<th></th>
<th>#Stocks</th>
<th>Mean</th>
<th>SD</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>#trades</td>
<td>2,157</td>
<td>2,015</td>
<td>4,798</td>
<td>79</td>
<td>432</td>
<td>1,797</td>
</tr>
<tr>
<td>dollar volume</td>
<td>2,157</td>
<td>0.025</td>
<td>0.062</td>
<td>0.001</td>
<td>0.006</td>
<td>0.021</td>
</tr>
<tr>
<td>1-min return</td>
<td>2,157</td>
<td>−0.007</td>
<td>0.255</td>
<td>−0.023</td>
<td>−0.001</td>
<td>0.011</td>
</tr>
<tr>
<td>1-min oib#</td>
<td>2,157</td>
<td>0.067</td>
<td>0.138</td>
<td>0.001</td>
<td>0.023</td>
<td>0.094</td>
</tr>
<tr>
<td>1-min oib$</td>
<td>2,157</td>
<td>3,052</td>
<td>7,077</td>
<td>4</td>
<td>520</td>
<td>2,616</td>
</tr>
</tbody>
</table>

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 billions USD (dollar volume), average one-minute mid-quote returns within the day in basis points (1-min return), average difference between the total number of trades initiated 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 USD) 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.

returns for a sample of 3,000 U.S. stocks from 1996 to 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 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 for 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 twenty 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 and 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...
Intraday return predictability regressions and summary statistics of stock-level efficiency measures

Panel A: Intraday return predictability regressions (with dependent variable $\text{Return}_{t}$)

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

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

<table>
<thead>
<tr>
<th>#Stocks</th>
<th>Mean</th>
<th>SD</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>OIB predictability</td>
<td>1,924</td>
<td>13</td>
<td>3.324</td>
<td>34.030</td>
<td>1.183</td>
</tr>
<tr>
<td>allquotes</td>
<td>1,924</td>
<td>8.714</td>
<td>24.971</td>
<td>0.598</td>
<td>1.538</td>
</tr>
<tr>
<td>2minutes</td>
<td>1,956</td>
<td>12.602</td>
<td>29.267</td>
<td>1.062</td>
<td>3.140</td>
</tr>
<tr>
<td>oib#</td>
<td>1,919</td>
<td>1.652</td>
<td>2.000</td>
<td>0.675</td>
<td>0.973</td>
</tr>
<tr>
<td>autocorrelation</td>
<td>1,837</td>
<td>0.023</td>
<td>0.033</td>
<td>0.003</td>
<td>0.023</td>
</tr>
<tr>
<td>Variance ratio</td>
<td>2,130</td>
<td>0.867</td>
<td>0.378</td>
<td>0.632</td>
<td>0.762</td>
</tr>
<tr>
<td>Hasbrouck</td>
<td>1,769</td>
<td>0.394</td>
<td>0.443</td>
<td>0.153</td>
<td>0.245</td>
</tr>
<tr>
<td>Put-call parity</td>
<td>1,535</td>
<td>2.581</td>
<td>2.043</td>
<td>1.467</td>
<td>2.091</td>
</tr>
</tbody>
</table>

(continued)

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).9 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 of 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 fifteen days over the sample period.

---

9 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.

10 We note that the literature (e.g., Boehmer and Wei 2003, Andrade, Chang, and Seasholes 2008) presents overwhelming evidence that intraday returns are predicted positively by lagged order flow and lagged returns.
The loadings on the first principal component almost always have the same

Table 2

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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Predictability</th>
<th>Variance ratio</th>
<th>Hasbrouck</th>
<th>Put-call parity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictability</td>
<td>100.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance ratio</td>
<td>(3.22)</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hasbrouck</td>
<td>35.95***</td>
<td>21.91***</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Put-call parity</td>
<td>25.63***</td>
<td>22.43***</td>
<td>52.69***</td>
<td>100.00</td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A of this table reports the average results of the return predictability regressions from Equation (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 $10^8$ 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 thirty times the ratio of the variance estimated from one-minute mid-quote returns to the variance estimated from thirty-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 1.1 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. *, **, and *** indicate significance at the 10%, 5%, and 1% level, 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.

To compress the five return predictability measures in Table 2 into a single measure to be used in the remainder of the paper, we take, for each stock, 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

Consistent with this literature, the vast majority of the estimated slope coefficients in panel A of Table 2 are positive. However, (large) negative slope coefficients in Equation (1) for a particular stock-day could arguably also be interpreted as evidence of inefficiency. Our main results are not materially affected when we take the absolute slope coefficient from Equation (1) as our stock-day return predictability measure or when we set negative stock-day coefficients to zero.

1165
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 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. This indicates that the resultant 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 this number is based on 1-to-20 days variance ratios (instead of on 1-to-30 minutes variance ratios as in our paper) and on a sample of NYSE stocks that is about half the size of our sample and likely tilted toward 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 I 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 I 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).
The Dynamics of Market Efficiency

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. Comovement in Efficiency Measures

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

3.1 Comovement in efficiency across stocks

To estimate the degree of comovement 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 comovement in efficiency for each stock \( i \) over the whole sample period in the following regression:

\[
E_{i,d} = \alpha_i + \beta_i MktEff_{i,d} + \gamma_i MktEff_{i,d-1} + \delta_i MktEff_{i,d+1} + \eta_{i,d},
\]

(2)

where \( E_{i,d} \) 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 Equation (2) for each stock with at least fifteen 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 comovement 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 Equation (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 comovement regressions in Equation (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.
This table reports the average results of the efficiency comovement regressions from Equation (2), estimated for each stock’s proportional quoted spread (PQSPR) on day \( t \). The independent variable \( MktEff_{d,i} \) 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 comovement 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 comovement 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 comovement 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 comovement 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.

Table 3 presents the results of our regressions to estimate comovement in each of the four efficiency measures across individual stocks. The table reports the average regression coefficients across all comovement regressions estimated by stock for each efficiency measure. The number of stocks for which we can estimate Equation (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 comovement in efficiency across stocks. The average coefficient on contemporaneous market-wide

<table>
<thead>
<tr>
<th>Efficiency measure:</th>
<th>Predictability</th>
<th>Variance ratio</th>
<th>Hasbrouck</th>
<th>Put-call parity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( MktEff_{d,i} )</td>
<td>0.483</td>
<td>0.699</td>
<td>1.528</td>
<td>1.064</td>
</tr>
<tr>
<td>( t )-stat avg</td>
<td>5.981</td>
<td>5.482</td>
<td>4.645</td>
<td>4.835</td>
</tr>
<tr>
<td>% positive</td>
<td>87.58</td>
<td>92.21</td>
<td>88.54</td>
<td>90.23</td>
</tr>
<tr>
<td>% + significant</td>
<td>69.30</td>
<td>72.46</td>
<td>66.00</td>
<td>68.44</td>
</tr>
<tr>
<td>( MktEff_{d,i-1} )</td>
<td>0.033</td>
<td>0.080</td>
<td>0.147</td>
<td>0.176</td>
</tr>
<tr>
<td>( t )-stat avg</td>
<td>0.534</td>
<td>0.123</td>
<td>0.309</td>
<td>0.844</td>
</tr>
<tr>
<td>% positive</td>
<td>57.35</td>
<td>51.40</td>
<td>56.86</td>
<td>61.20</td>
</tr>
<tr>
<td>% + significant</td>
<td>27.71</td>
<td>11.66</td>
<td>18.53</td>
<td>28.37</td>
</tr>
<tr>
<td>( MktEff_{d,i+1} )</td>
<td>0.033</td>
<td>0.098</td>
<td>-0.061</td>
<td>0.180</td>
</tr>
<tr>
<td>( t )-stat avg</td>
<td>0.712</td>
<td>0.064</td>
<td>0.313</td>
<td>0.834</td>
</tr>
<tr>
<td>% positive</td>
<td>58.76</td>
<td>50.81</td>
<td>59.30</td>
<td>62.06</td>
</tr>
<tr>
<td>% + significant</td>
<td>30.70</td>
<td>10.44</td>
<td>19.10</td>
<td>28.64</td>
</tr>
<tr>
<td>( R^2 ) (in %)</td>
<td>6.47</td>
<td>3.23</td>
<td>6.73</td>
<td>8.81</td>
</tr>
<tr>
<td>adj ( R^2 ) (in %)</td>
<td>5.21</td>
<td>2.31</td>
<td>4.94</td>
<td>7.96</td>
</tr>
<tr>
<td># regressions</td>
<td>1,707</td>
<td>2,041</td>
<td>1,597</td>
<td>1,505</td>
</tr>
</tbody>
</table>

This table reports the average results of the efficiency comovement regressions from Equation (2), estimated for each NYSE stock in the sample based on daily data over the whole sample period. The dependent variable \( E_{f,i,d} \) is the efficiency of stock \( i \) on day \( d \), orthogonalized with respect to stock \( i \)’s proportional quoted spread (PQSPR) on day \( d \). 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 comovement in each of the four efficiency measures across individual stocks. The table reports the average regression coefficients across all comovement regressions estimated by stock for each efficiency measure. The number of stocks for which we can estimate Equation (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 comovement 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. This 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). In line with their recommendation, for each column in Table 3 we compute the average pairwise correlation ($\rho$) between the residuals across all $N$ regressions and then multiply the standard errors by $[1+(N-1)\rho]^{1/2}$. The resultant adjusted $t$-statistics also indicate statistically significant comovement 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 comovement regressions in Equation (2) can be interpreted as an alternative way to gauge the degree of efficiency comovement across stocks. The average (adjusted) $R^2$'s of the comovement 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 comovement 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 Equation (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 (2013). For similar monthly regressions to estimate comovement in liquidity for NYSE stocks over the period 1995-2009. We also note that portfolios of stocks might alleviate estimation noise and expose a stronger image of comovement. Accordingly, we revisit the comovement analysis in Table 3 by running regressions of the value-weighted average efficiency of the stocks in each of the five industries defined on Kenneth French’s Web site 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 comovement regressions in Equation (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 comovement in efficiency across stocks, suggesting that common economic forces may drive variation in the price efficiency of individual securities.

### 3.2 Comovement in efficiency across measures

We now turn to the question of 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 and 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.

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 some clear common patterns across several of the measures, but also some idiosyncratic fluctuations. For example, *Predictability*, *Hasbrouck*, and *Put-call parity* all show a long-run downward-sloping trend that 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 in *Variance ratio*, seems largely independent of time variation in the other measures.

In Table 4, we assess the extent of comovement across the five different monthly market-wide efficiency measures by presenting Spearman rank correlations across all of the measures (Pearson correlations are similar).

---

11 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 were able to closely match their results. In particular, they reported an average slope coefficient on thirty minute returns that are lagged 13 intervals (or one full trading day) of 1.19 (their Table I), while we obtained an estimate of 1.15 over their sample period, 2001-2005.
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.

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

<table>
<thead>
<tr>
<th></th>
<th>Predictability</th>
<th>Variance ratio</th>
<th>Hasbrouck</th>
<th>Put-call parity</th>
<th>HKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictability</td>
<td>100.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance ratio</td>
<td>−36.64***</td>
<td>−35.78***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hasbrouck</td>
<td>59.75***</td>
<td>42.07***</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Put-call parity</td>
<td>18.83**</td>
<td>−37.10***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HKS</td>
<td>−3.49</td>
<td>−3.58</td>
<td>14.87**</td>
<td>−3.66</td>
<td>100.00</td>
</tr>
<tr>
<td>(0.64)</td>
<td>(0.64)</td>
<td>(0.05)</td>
<td>(0.63)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 thirty-minute stock returns on returns over the same interval on the previous day (averaged over all thirty-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. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

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 (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 on more than a single measure of market efficiency in empirical research.

12 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) did not detrend the underlying variables in their analysis of the extent of covariation across high- and low-frequency measures of market liquidity.

1172
4. What Drives Systematic Variation in Market Efficiency?

Our results in Section 3 indicate that market efficiency measures exhibit significant comovement across individual stocks and with each other. These findings suggest 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 a 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; doing so can lead to greater deviations from efficient pricing and potentially trigger feedback loops that cause inefficiencies and funding constraints to worsen in lockstep. 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 Treasury-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 billions USD. (Hedge fund AUM) in our analysis. We expect greater assets under

---

13 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 (2014).
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 \textit{Hedge fund AUM} is associated with an improvement in systematic market efficiency.

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. This evidence 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, thereby 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 \((\text{Quotes/Volume})\). Our third hypothesis (H3) is that a positive shock to \textit{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 \textit{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 \textit{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 subgroups 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 \textit{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, \textit{Hedge fund AUM}, \textit{Quotes/Volume}, and \textit{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: \textit{Predictability}, Hasbrouck, Put-call parity, and HKS. We exclude the \textit{Variance ratio} from the principal component analysis because of its negative correlation with the other measures. However, our results are robust to including the \textit{Variance ratio} measure. We follow Hasbrouck and Seppä (2000) and extract the principal components based on the correlation matrix. We find that the first
The Dynamics of Market Efficiency

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. 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 measure. 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 resultant 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 subgroups 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 separately obtain a systematic market efficiency component for the low and high turnover portfolios 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.
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.

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 for 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). 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,

14 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.
### Table 5
Summary statistics of potential determinants of systematic variation in efficiency

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

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 billions USD, obtained from Matti Suominen and LIPPER-TASS (see Jylhä, Rinne, and Suominen 2014). **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).

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 billions USD, obtained from Matti Suominen and LIPPER-TASS (see Jylhä, Rinne, and Suominen 2014). 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).

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

Panel A: All stocks

<table>
<thead>
<tr>
<th></th>
<th>AEFF</th>
<th>Volatility</th>
<th>Quotes/Volume</th>
<th>Hedge fund AUM</th>
<th>TED spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEFF</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>44.83***</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Quotes/Volume</td>
<td>−2.51</td>
<td>59.07***</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hedge fund AUM</td>
<td>(0.74)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>TED spread</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

Panel B: Low turnover stocks

<table>
<thead>
<tr>
<th></th>
<th>AEFF</th>
<th>Volatility</th>
<th>Quotes/Volume</th>
<th>Hedge fund AUM</th>
<th>TED spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEFF</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>52.03***</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Quotes/Volume</td>
<td>7.55</td>
<td>47.91***</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hedge fund AUM</td>
<td>(0.32)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>TED spread</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

Panel C: High turnover stocks

<table>
<thead>
<tr>
<th></th>
<th>AEFF</th>
<th>Volatility</th>
<th>Quotes/Volume</th>
<th>Hedge fund AUM</th>
<th>TED spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEFF</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>50.98***</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Quotes/Volume</td>
<td>−6.78</td>
<td>46.38***</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hedge fund AUM</td>
<td>(0.37)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>TED spread</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

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 3 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. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

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.
The Dynamics of Market Efficiency

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).

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

15 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.
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.
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, nontrivial 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 other endogenous variables.
We also estimated four separate VARs with each of the individual efficiency measures underlying \( \text{AEFF} \) as key endogenous variables. 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 \( \text{Quotes/Volume} \) is most pronounced for the \( \text{Predictability} \) measure, while the negative effect of \( \text{Hedge fund AUM} \) is strongest for \( \text{Hasbrock and HRS} \). These findings could be indicative of different types of arbitrageurs focusing on different

---

Table 7

<table>
<thead>
<tr>
<th>VARs of systematic market efficiency (( \text{AEFF} ))</th>
<th>Panel A: All stocks</th>
<th>Panel B: Low turnover stocks</th>
<th>Panel C: High turnover stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{AEFF}_m )</td>
<td>0.610**</td>
<td>0.685**</td>
<td>0.546**</td>
</tr>
<tr>
<td>( \text{AEFF}_{m-1} )</td>
<td>0.979**</td>
<td>1.069**</td>
<td>0.929**</td>
</tr>
<tr>
<td>( \text{AEFF}_{m-2} )</td>
<td>1.204**</td>
<td>1.171**</td>
<td>1.204**</td>
</tr>
<tr>
<td>( \text{Volatility}_m )</td>
<td>1.344**</td>
<td>1.259**</td>
<td>1.402**</td>
</tr>
<tr>
<td>( \text{Volatility}_{m-1} )</td>
<td>0.349**</td>
<td>0.369**</td>
<td>0.344**</td>
</tr>
<tr>
<td>( \text{Volatility}_{m-2} )</td>
<td>0.616**</td>
<td>0.534**</td>
<td>0.589**</td>
</tr>
<tr>
<td>( \text{Volatility}_{m-3} )</td>
<td>0.809**</td>
<td>0.623**</td>
<td>0.759**</td>
</tr>
<tr>
<td>( \text{Quotes/Volume}_m )</td>
<td>-0.118**</td>
<td>-0.017</td>
<td>-0.127**</td>
</tr>
<tr>
<td>( \text{Quotes/Volume}_{m-1} )</td>
<td>-0.211**</td>
<td>0.017</td>
<td>-0.268**</td>
</tr>
<tr>
<td>( \text{Quotes/Volume}_{m-2} )</td>
<td>-0.289**</td>
<td>0.065</td>
<td>-0.402**</td>
</tr>
<tr>
<td>( \text{Quotes/Volume}_{m-3} )</td>
<td>-0.355*</td>
<td>0.110</td>
<td>-0.519**</td>
</tr>
<tr>
<td>( \text{Hedge fund AUM}_m )</td>
<td>-0.150</td>
<td>-0.106</td>
<td>-0.226**</td>
</tr>
<tr>
<td>( \text{Hedge fund AUM}_{m-1} )</td>
<td>-0.235*</td>
<td>-0.148</td>
<td>-0.380**</td>
</tr>
<tr>
<td>( \text{Hedge fund AUM}_{m-2} )</td>
<td>-0.278</td>
<td>-0.161</td>
<td>-0.485**</td>
</tr>
<tr>
<td>( \text{Hedge fund AUM}_{m-3} )</td>
<td>-0.291</td>
<td>-0.157</td>
<td>-0.554**</td>
</tr>
<tr>
<td>( \text{Ted spread}_m )</td>
<td>0.172**</td>
<td>0.256**</td>
<td>0.102*</td>
</tr>
<tr>
<td>( \text{Ted spread}_{m-1} )</td>
<td>0.283**</td>
<td>0.450**</td>
<td>0.180*</td>
</tr>
<tr>
<td>( \text{Ted spread}_{m-2} )</td>
<td>0.358*</td>
<td>0.600**</td>
<td>0.247*</td>
</tr>
<tr>
<td>( \text{Ted spread}_{m-3} )</td>
<td>0.406*</td>
<td>0.642**</td>
<td>0.285</td>
</tr>
</tbody>
</table>

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, \( \text{Quotes/Volume} \), \( \text{Volatility} \), and the systematic component of market efficiency \( \text{AEFF} \) (see variable descriptions in Tables 5-6 and Figure 4). The endogenous variables \( \text{AEFF} \), \( \text{Quotes/Volume} \), and \( \text{Volatility} \) are each constructed separately for each panel based on the relevant sample of stocks. Each panel shows the cumulative response of \( \text{AEFF} \) 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 4, *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

The negative effect of a shock to \( \text{Quotes/Volume} \) on \( \text{AEFF} \) in Figure 4 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 \( \text{AEFF} \) 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.
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 measures vary over time, whether they comove 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 comove 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 (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 from these daily stock-level measures, as well as from 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 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 and Table, possibly as a result of the somewhat greater degree of noise in each of the individual measures.
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 comovement 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 to fixed income securities, foreign exchange, and derivatives. These and other issues are left for future research.

References


The Dynamics of Market Efficiency


The Dynamics of Market Efficiency


