Illiquidity and the Measurement of Stock Price Synchronicity

Gassen, J.; Skaife, H.A.; Veenman, D.

DOI
10.1111/1911-3846.12519

Publication date
2020

Document Version
Final published version

Published in
Contemporary Accounting Research

Illiquidity and the Measurement of Stock Price Synchronicity*

JOACHIM GASSEN, *Humboldt University Berlin*

HOLLIS A. SKAIFE, *University of California – Davis*

DAVID VEENMAN, *University of Amsterdam*†

ABSTRACT

This paper demonstrates that measures of stock price synchronicity based on market model R^2s are predictably biased downward as a result of stock illiquidity, and that previously employed remedies to correct market model betas for measurement bias do not fix R^2. Using a large international sample of firm-years, we find strong negative and nonlinear relations between illiquidity and R^2 across countries, across firms, and over time. Because variables of interest frequently relate to illiquidity as well, we illustrate the consequences of not controlling for illiquidity in synchronicity research. More generally, we demonstrate the importance of using nonlinear control variable methods. Overall, we conclude that the illiquidity-driven measurement bias in R^2 provides an explanation for why prior research finds low-R^2 firms to have weak information environments, and suggest future research carefully evaluate the sensitivity of its results to nonlinear controls for illiquidity.

Illiquidité et mesure de la synchronicité du cours des actions

RÉSUMÉ

Les auteurs démontrent que les mesures de la synchronicité du cours des actions basées sur les R^2 du modèle de marché sont biaisées vers le bas en raison de l'illiquidité des actions, et que les solutions utilisées jusqu'à maintenant pour corriger les coefficients bêta du modèle de marché de manière à tenir compte du biais de ces mesures n'ont pas résolu la question du coefficient R^2. L'étude d'un vaste échantillon international d'observations société-année à laquelle procèdent les auteurs les amène à constater l'existence de fortes relations négatives et non linéaires entre l'illiquidité et le coefficient R^2 dans l'ensemble des pays, dans l'ensemble des sociétés et dans le temps. Compte tenu du fait que les variables critiques affichent souvent aussi un lien avec l'illiquidité, les auteurs illustrent les conséquences de l'absence de contrôle du facteur d'illiquidité dans la recherche sur la synchronicité. De manière plus générale, ils montrent l'importance du recours à des méthodes faisant appel à une variable de contrôle non linéaire. Globalement, les auteurs concluent que le biais de la mesure du coefficient R^2 induit par l'illiquidité contribue à expliquer pourquoi les études antérieures ont indiqué que les sociétés dont le R^2

* Accepted by Dan Segal. This paper supersedes an earlier working paper titled “Does Stock Price Synchronicity Represent Firm-Specific Information?: International Evidence.” We are grateful to Ryan LaFond for his contributions to this earlier project, which can be found at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=768024. We thank two anonymous reviewers, Dan Segal, Michael Welker, Sanjay Bissessur, Peter Easton, Frank Ecker, Igor Goncharov, Katharina Hombach, Felix Lamp, Stefan Obernberger, Helen Popper, Bill Rees, Sandra Schafhautle, participants at the European Accounting Association annual meeting in Glasgow, and workshop participants at LMU Munich and the Frankfurt School of Finance and Management for helpful comments on the current version of the paper. Joachim Gassen acknowledges financial support by the Deutsche Forschungsgemeinschaft (German research foundation, DFG) under Project-ID 403041268 – TRR 266. We thank I/B/E/S for providing data on analyst coverage and earnings announcement dates.

† Corresponding author.

Contemporary Accounting Research Vol. 37 No. 1 (Spring 2020) pp. 419–456
Contemporary Accounting Research published by Wiley Periodicals, Inc. on behalf of CAAA.
doi:10.1111/1911-3846.12519

était faible avaient des environnements pauvres en information; ils suggèrent que soit évaluée avec soin dans les recherches à venir la sensibilité de ces résultats aux contrôles non linéaires de l'illiquidité.

1. Introduction

The dissemination and acquisition of information play a key role in the price formation process in equity markets (e.g., Grossman and Stiglitz 1976, 1980; Verrecchia 1982; Easley and O'Hara 2004; Kelly and Ljungqvist 2012). In this regard, prior research frequently studies the efficiency with which stock prices incorporate firm-specific information. Following Roll (1988) and Morck et al. (2000), a large literature examines the informational efficiency of prices, or stock price informativeness, by studying measures of stock price synchronicity based on market model R^2 values. The premise underlying the use of R^2 is the assumption that if market returns explain relatively more of the variation in firms' stock returns, relatively less firm-specific information is impounded in price and firms' stock prices are therefore less informationally efficient.

Based on this assumption, several studies use synchronicity metrics to examine the effects of corporate transparency or information intermediaries on the mix of information in stock prices.[1] Conceptually, corporate transparency can lead to lower synchronicity in several ways.[2] First, for firms that are more transparent, the amount of *public* firm-specific information available to investors on a given day increases relative to the amount of market-wide information (e.g., Hutton et al. 2009). Second, to the extent that greater transparency reduces adverse selection, the costs of acquiring and trading on *private* firm-specific information go down. The resulting increase in informed trading increases the relative amount of firm-specific information that is reflected in prices (e.g., Grossman and Stiglitz 1980; Glosten and Milgrom 1985; Kyle 1985; Veldkamp 2006). Third, Jin and Myers (2006) argue that insiders of less transparent firms are able to asymmetrically capture part of firm-specific risk, which makes the stock price relatively less (more) sensitive to firm-specific (market-wide) news.

Despite the extensive use of R^2 measures in the literature, a significant debate exists on whether low R^2 values are more representative of firm-specific information or simply noise in prices. Consistent with the noise argument, recent studies empirically find that low R^2s are associated with weaker information environments (e.g., Teoh et al. 2009; Chan and Chan 2014; Kelly 2014; Li et al. 2014). Our study informs this debate by demonstrating the consequences of measurement bias in R^2 induced by the lack of liquidity in a firm's stock ("illiquidity"), an issue that tends to be associated with weaker information environments.

The measurement problem we focus on is straightforward. When there is little or no market for a firm's stock, the stock trades infrequently and is illiquid. Prior research demonstrates how non-trading biases the measurement of systematic risk (beta) factors (e.g., Scholes and Williams 1977; Dimson 1979; Lo and MacKinlay 1990). We extend this literature by modeling how illiquidity affects market model R^2s. The underlying idea is that, ceteris paribus, the observed returns of an illiquid stock are less likely to co-move with the market because illiquidity hinders the stock price from moving when market-wide information changes the value of the firm. This affects both market model regression coefficients and R^2s. Using zero daily returns as an outcome of illiquidity (Lesmond et al. 1999), we demonstrate econometrically how illiquidity causes a downward bias in R^2.[3]

1. Several studies link synchronicity to factors related to corporate transparency, such as Jin and Myers (2006), Haggard et al. (2008), Hutton et al. (2009), Gul et al. (2010), Gul et al. (2011), Bartram et al. (2012), Armstrong et al. (2012), Kim et al. (2012), Li et al. (2014), Peterson et al. (2015), Piotroski et al. (2015), Dong et al. (2016), Grewal et al. (2017), and Choi et al. (2019). Other studies such as Bissessur and Hodgson (2012), Kim and Shi (2012), Wang and Yu (2015), and Barth et al. (2018) link changes in accounting standards to changes in synchronicity, while Piotroski and Roulstone (2004), Chan and Hameed (2006), Crawford et al. (2012), Muslu et al. (2014), and Wang (2019) relate sell-side analysts' research to the relative amount of firm-specific information reflected in prices.
2. In section 2, we discuss how these conceptual links between corporate transparency and R^2 are more ambiguous in multiperiod settings. We thank an anonymous reviewer for raising this important point.
3. With bias in R^2, we refer to the difference between the observed R^2 from market model regressions and the unobservable R^2 that would be obtained if firm-level stock returns were not affected by non-trading effects.

To document the empirical implications of our argument, we empirically examine a large sample of listed firms from 50 countries. To first assess whether R^2 measures are able to capture firm-specific information in prices in our data, we initially examine the extent to which large flows of firm-specific information actually manifest in lower R^2. Using annual earnings announcements, we find that firm-specific information flows *are* associated with statistically significant reductions in R^2, but these effects are economically modest (as in, e.g., Roll 1988). Partitioning the data by firms' stock illiquidity, we find that these effects are larger for liquid firms, while measurement bias masks these effects for the most illiquid firms. Hence, these analyses suggest that researchers can use R^2 measures to identify the relative amount of firm-specific information in prices, but that measurement bias makes low R^2 appear to reflect more noise in prices.

We next find that countries with low R^2s tend to exhibit high illiquidity, as observed through a strong correlation of −0.67 between average country-level R^2 and average country-level zero-return frequency. Focusing on temporal variation, we also find that increases and decreases in illiquidity coincide with decreases and increases in R^2, respectively. For instance, the low levels of R^2 in the 1990s in the United States and the sharp increase in R^2 in the 2000s (e.g., Campbell et al. 2001; Morck et al. 2013) relate to relatively high and low levels of illiquidity over the same periods, respectively. Next, we document that, within countries, illiquidity and R^2 are consistently negatively related across firms, and find that the empirical links between illiquidity and R^2 we uncover are nonlinear.

Following the prior literature on non-trading effects (e.g., Scholes and Williams 1977; Dimson 1979), the synchronicity literature commonly adds lags and leads of market returns to the market model, or reduces the return frequency (e.g., weekly instead of daily returns), to address measurement problems. However, these approaches are designed to correct the measurement of beta and not to fix R^2. Specifically, while these approaches solve the errors-in-variables problem in the market model (i.e., the explanatory variables are measured with error), which leads to consistent estimates of beta, they do not eliminate the bias in R^2. We explain and demonstrate empirically that these common remedies reduce the measurement bias in R^2 only to a limited extent.[4]

The result of a strong negative relation between illiquidity and R^2 suggests researchers should be cautious when interpreting synchronicity metrics. Variables of interest are often directly or indirectly related to illiquidity, which implies that measurement bias can confound inferences from tests of their relation with R^2. For example, if a variable of interest has no relation with synchronicity, but is negatively (positively) related to illiquidity, the measurement bias in R^2 leads to an omitted variable problem and the observation of a significant positive (negative) relation between the variable of interest and R^2 when illiquidity is not controlled for. If the variable of interest does have a negative (positive) relation with synchronicity and is itself negatively (positively) related to illiquidity, the observed relation between the variable of interest and R^2 is biased in the opposite direction when illiquidity is not controlled for. This omitted variable problem limits the researcher's ability to identify the relation of interest.

With "controlling for," we refer to any parametric or nonparametric approach that makes the statistical relation of interest conditional on a (set of) control variable(s) (Morgan and Winship 2015; Gow et al. 2016).[5] Because the link between illiquidity and R^2 is nonlinear, we analyze the consequences of using flexible control variable approaches that allow the effects of a control variable to be nonlinear. Specifically, while traditional control variable approaches rely on the assumption of a linear functional form, we focus on a fixed effects or "least squares

4. We return to this discussion in section 2. In Appendix 1, our simulations further demonstrate that these common remedies do not eliminate the bias in R^2.

5. Examples of such approaches are matching methods, or regressions in which control variables are split up into fixed effects (Angrist and Pischke 2009). At the same time, we recognize the limitations of control variable approaches, since research designs should typically not include control variables that are part of the conceptual mechanism investigated ("mediators") or variables that are outcome variables ("colliders"). In Appendix 2, we discuss these issues and demonstrate the importance to researchers of using causal diagrams to better understand whether controlling for a variable helps or hinders them in identifying the theoretically predicted relations.

dummy variable" (LSDV) approach (e.g., Gormley and Matsa 2014). Here, the control variable of interest is split into n dummy variables that each take on a value of one for specific values of the control variable (zero otherwise). This approach allows for a more flexible functional form of the influence of a confounding factor (e.g., Angrist and Pischke 2009) and picks up more of the relevant variation in synchronicity than a traditional (linear) control variable approach.

We illustrate these issues for the relation between R^2 and analyst coverage and corporate transparency, respectively—two variables related to liquidity (e.g., Alford and Berger 1999; Leuz and Verrecchia 2000). Consistent with prior research, analyst coverage is positively associated with synchronicity in our sample (Piotroski and Roulstone 2004; Chan and Hameed 2006). However, after controlling for illiquidity using the LSDV approach, the positive relation between analyst coverage and R^2 largely disappears. This result suggests that conclusions about the role of analysts in shaping the mix of information in stock prices are sensitive to controls for illiquidity. Next, using country-level measures of transparency from Jin and Myers (2006) and Bushman et al. (2004), we identify, based on our data, significant negative relations between transparency and R^2 *only* after illiquidity is controlled for. This result suggests that failing to control for illiquidity can also limit researchers' ability to identify relations of interest.

This paper makes the following contributions. First, our research informs the debate about the use of stock price synchronicity as a measure of price informativeness.[6] The synchronicity literature hinges on the validity of the premise that low R^2 values are associated with more informative stock prices. However, recent studies by Teoh et al. (2009), Chan and Chan (2014), Li et al. (2014), and Kelly (2014) find that low R^2 is associated with variables that indicate less informative stock prices and conclude that low R^2 measures noise. Our study demonstrates *why* low R^2 tends to be associated with weaker information environments, by showing how measurement bias induces a strong negative relation between R^2 and illiquidity.[7] We also find that while firm-specific information flows do manifest in lower R^2, measurement bias masks this effect for less liquid stocks, thereby making low R^2 *appear* to reflect more noise in prices.

Second, our study informs prior and future research that examines synchronicity in settings in which variables of interest potentially relate to illiquidity.[8] Because previously employed adjustments to market model estimations do not eliminate the bias in R^2 metrics, we argue that future synchronicity research should explicitly control for illiquidity. We empirically illustrate how controlling for illiquidity in synchronicity research can both (i) eliminate a previously identified significant relation of interest and (ii) help researchers identify significant effects where they exist. In doing so, we also contribute to the literature more broadly by demonstrating the implications of using nonlinear control variable methods instead of the traditional (linear) control variable approaches typically used in the extant literature.

Lastly, this paper contributes to the literature that examines the effects of sell-side analysts' research on stock price informativeness. While prior research has examined these effects for emerging markets (Chan and Hameed 2006) or the United States (Piotroski and Roulstone 2004), we extend this literature by examining the relation between analyst coverage and synchronicity for a comprehensive sample of firms from 50 countries. Importantly, we demonstrate that

6. See, for example, Durnev et al. (2003), Irvine and Pontiff (2009), Teoh et al. (2009), Xing and Anderson (2011), Bartram et al. (2012), Kelly (2014), and Li et al. (2014).

7. Chan et al. (2013) also relate synchronicity to illiquidity. They predict that synchronicity affects liquidity based on the assumption that investors are more likely to transact in stocks that co-move more with market fundamentals and, similar to us, document a robust positive correlation between stock price synchronicity and liquidity. Our study is fundamentally different from Chan et al. (2013) since we document a strong measurement-based effect of illiquidity on stock price synchronicity, supporting a causal interpretation that runs opposite to that of Chan et al. (2013).

8. Examples of such settings are the adoption of new accounting standards or changes in regulatory enforcement (Christensen et al. 2013), international differences in disclosure transparency (Lang et al. 2012), analysts' coverage decisions (Roulstone 2003), the quality of financial reporting (Bhattacharya et al. 2013), and management forecasts (Coller and Yohn 1997), all of which have previously been linked to stock price informativeness.

conclusions of a positive relation between synchronicity and analyst coverage are sensitive to controls for illiquidity. Controlling for illiquidity is important in this setting, because the argument that analysts increase the amount of market-wide information in prices rests on the assumed mechanism of analysts directly providing this information to market participants. Thus, it does not rely on changes in illiquidity as a channel through which the information is impounded in prices. After controlling for illiquidity, we do not find consistent evidence to suggest that analysts increase the relative amount of market-wide information impounded in stock prices.

2. Firm-specific information, illiquidity, and the measurement of stock price synchronicity

Firm-specific information and stock price synchronicity

Consider the following market model as a simplified representation of the relative change in value (R^*_{jt}) of firm j's stock in period t (e.g., Dimson 1979; Lo and MacKinlay 1990; Lesmond et al. 1999):

$$R^*_{jt} = \beta_j R_{mt} + \varepsilon_{jt}. \tag{1}$$

This model depicts R^*_{jt} as a function of both common, market-wide information ($\beta_j R_{mt}$), and firm-specific, idiosyncratic news (ε_{jt}). In a frictionless market without trading costs, public information is impounded in price immediately and the *observed* return of a firm (R_{jt}) should equal the *true* underlying return R^*_{jt} (e.g., Campbell et al. 1997). Researchers typically estimate this model using OLS regression for a set of firm-specific return observations over a finite window (e.g., calendar year) and fixed interval (e.g., daily). Coefficient β_j ($= Cov\left[R^*_{jt}, R_{mt}\right] / Var[R_{mt}]$) then captures the sensitivity of returns to non-diversifiable common information. The explanatory power (R^2) from this regression is frequently labeled stock price synchronicity and used as an inverse measure of a firm's stock price informativeness.

Conceptually, low R^2 can be interpreted as suggesting that variation in a firm's stock returns is driven relatively more by firm-specific information as opposed to common information (Roll 1988; Morck et al. 2000). Therefore, research on stock price synchronicity could provide important insights into the effects of, for example, corporate transparency on price informativeness. For example, when a firm is completely opaque, investors cannot observe the firm-specific component of changes in value and, as a result, the firm's stock price would move only with common information. When the firm becomes more transparent, relatively more of the firm-specific information is revealed and becomes public in each trading period, which leads to relatively greater firm-specific return variation and lower R^2 values (e.g., Hutton et al. 2009).[9]

Corporate transparency can also affect R^2 indirectly. If corporate transparency reduces adverse selection, the costs of acquiring and trading on private firm-specific information decline. The increase in informed trading that results from these reduced costs can increase the relative amount of firm-specific information that is reflected in prices (e.g., Grossman and Stiglitz 1980). Morck et al. (2000) similarly argue that investor protection increases the benefits to arbitrage and investors' willingness to generate, and trade on, private firm-specific information. Jin and Myers (2006) suggest an alternative, agency-cost related mechanism, by arguing that corporate transparency affects the extent to which managers can extract private rents.[10]

9. Hutton et al. (2009, 67) state that in studies linking transparency to R^2, the "common point of departure is the notion that greater transparency and more complete revelation of firm-specific information should reduce R^2."

10. Jin and Myers (2006) predict that the level of opacity of firms affects the degree to which the market versus insider managers bear the idiosyncratic risk of the firm. In their model, opacity helps managers extract part of the firm's cash flows as private rents in case of good performance, while they give up part of their private wealth to compensate in case of bad performance. Managers' actions therefore dampen the variation in cash flows, which dampens the variation in idiosyncratic information that is impounded in market prices. This, in turn, increases R^2, and this effect strengthens with firm opacity.

Another example of a setting in which R^2 is conceptually useful is in tests of the role of information intermediaries, such as sell-side analysts, in shaping the relative mix of information reflected in prices (Piotroski and Roulstone 2004; Chan and Hameed 2006; Crawford et al. 2012). If analysts' research generates publicly available firm-specific information, more of the variation in firms' stock returns should be explained by firm-specific information, thereby lowering R^2. Alternatively, if analysts' research relates primarily to helping the market dissect and link common information to a firm, such as economy-wide or industry-level information, these activities would enhance the efficiency of common information in aggregating information about the firm, thereby increasing R^2 (Piotroski and Roulstone 2004).

These predicted effects on R^2 are not unambiguous in a multiperiod setting. In the above examples, the dissemination and impounding of firm-specific information happen at a specific point in time, while in reality firm-specific information is often revealed, and thus impounded in price, with a delay. As can be shown analytically, the critical factor linking transparency to R^2 is the relative *degree* to which firm-specific information remains private on any given day due to firms' opacity.[11] If instead disclosure is *discrete*, in that all or none of the private information is disseminated on a day, the frequency of disclosure should have little effect on R^2 as all information is impounded in prices in multiple-day windows. For example, R^2 for a transparent firm fully disclosing its private information every day should not be different from an opaque firm that only discloses this information on the last day of each quarter.[12] Factors such as a firm's degree of opaqueness and analyst coverage, however, may affect the *degree* to which firm-specific information remains private on a given day, and therefore affect R^2.

Overall, we conclude that synchronicity metrics based on R^2 can conceptually be useful for studies of the effects of different disclosure and information mechanisms on the price formation process in equity markets. At the same time, our discussion suggests that it is essential for researchers to carefully assess whether and how the mechanisms they investigate should actually be expected to lead to variation in R^2.[13]

Econometric implications of stock illiquidity for synchronicity measurement

Besides the transparency-related arguments discussed above, the observed return R_{jt} for a firm can differ from its true return R^*_{jt} because of trading frictions. When securities trade infrequently, the price observed for a security is often the outcome of an informed trade that occurred earlier, and this infrequent trading has important implications for estimations of the market model (Dimson 1979; Lo and MacKinlay 1990; Scholes and Williams 1977). Non-trading of a security can occur because of a liquidity premium (e.g., Amihud and Mendelson 1986), and the marginal investor will not trade a security unless the value of new information is greater than this liquidity

11. Proofs and illustrations using simulations analyses are available from the authors upon request.
12. This example is comparable to Lo and MacKinlay's (1990) modeling of the effects of non-trading on return variance (see more on this in the next section). In their setting, non-trading leads to zero returns. In the setting described here, nondisclosure leads to zero *idiosyncratic* returns. Lo and MacKinlay (1990) show that zero returns have limited effects on return variance (see their equation 2.11). In the nondisclosure setting described here, zero idiosyncratic returns similarly have limited effect on overall return variance and, therefore, R^2.
13. In discussing Kim and Shi (2012), Christensen (2012, 523–524) provides a good example of such. He states that "stock price synchronicity may not be the ideal outcome variable if the objective is to identify a causal effect of IFRS adoption. Although it is straightforward to argue that IFRS could affect the quality of annual, and perhaps interim, reports, it is less clear why annual reports should affect stock price synchronicity. Kim and Shi measure stock price synchronicity weekly over the year, but annual reports are disclosed, by definition, only once a year. It is unclear why infrequent disclosures should affect information flow throughout the year." In a similar vein, Jin and Myers (2006) argue that when "hidden news is revealed after a stable lag . . . the average amount of firm-specific information released in any period is the same as for a transparent firm. Average firm-specific variance and R^2 are not affected by delayed reporting."

premium. Otherwise, a zero return is observed and the true return R_{jt}^* is not incorporated into the security's market price at time t.

As a result of zero returns due to non-trading, the continuously compounded return on a subsequent day *with* trading can be depicted as follows (Lo and MacKinlay 1990):

$$R_{jt} = \sum_{k_{jt}=0}^{K_{jt}} R_{jt-k}^* = R_{jt}^* + \sum_{k_{jt}=1}^{K_{jt}} R_{jt-k}^* = R_{jt}^* + \theta_{jt}, \tag{2}$$

where K_{jt} is the duration of non-trading prior to day t and θ_{jt} is the sum of true returns on the previous consecutive non-trading days. If we denote the probability of non-trading as p_{jt} and introduce a random variable δ_{jt} that takes a value of one for days without trading and zero for days with trading, we can decompose the observed return of security j for period t as follows:

$$R_{jt} = \begin{cases} 0 & \text{if } \delta_{jt} = 1 \text{ with probability } p_j \\ R_{jt}^* + \theta_{jt} & \text{if } \delta_{jt} = 0 \text{ with probability } \left(1 - p_j\right). \end{cases} \tag{3}$$

Using the law of iterated expectations, the covariance between observed returns (R_{jt}) and market returns (R_{mt}) can be rewritten as:

$$\begin{aligned} Cov\left[R_{jt}, R_{mt}\right] &= E\left[\left(R_{jt} - E\left[R_{jt}\right]\right)\left(R_{mt} - E[R_{mt}]\right)\right] = E\left[E\left[\left(R_{jt} - E\left[R_{jt}\right]\right)\left(R_{mt} - E[R_{mt}]\right) \mid \delta\right]\right] \\ &= E[\delta = 1] E[(0)(R_{mt} - E[R_{mt}])] + E[\delta = 0] E\left[\left(R_{jt}^* + \theta_{jt} - E\left[R_{jt}^* + \theta_{jt}\right]\right)\left(R_{mt} - E[R_{mt}]\right)\right] \\ &= \left(1 - p_j\right) E\left[\left(R_{jt}^* + \theta_{jt} - E\left[R_{jt}^* + \theta_{jt}\right]\right)\left(R_{mt} - E[R_{mt}]\right)\right] \\ &= \left(1 - p_j\right) E\left[\left(R_{jt}^* - E\left[R_{jt}^*\right]\right)\left(R_{mt} - E[R_{mt}]\right)\right] \\ &\quad + \left(1 - p_j\right) E\left[\left(\theta_{jt} - E\left[\theta_{jt}\right]\right)\left(R_{mt} - E[R_{mt}]\right)\right]. \end{aligned} \tag{4a}$$

Assuming that $Cov[\theta_{jt}, R_{mt}] = 0$, we obtain[14]

$$Cov\left[R_{jt}, R_{mt}\right] = \left(1 - p_j\right) Cov\left[R_{jt}^*, R_{mt}\right], \tag{4b}$$

that is, the covariance for observed returns ($Cov[R_{jt}, R_{mt}]$) equals a fraction $(1 - p_j)$ of the covariance for true returns ($Cov\left[R_{jt}^*, R_{mt}\right]$). Because only $Cov[R_{jt}, R_{mt}]$ is observed, non-trading results in downwardly biased estimates of β when holding constant the variance of returns on the market index: $b_j = q_j \beta_j$, where $q_j = (1 - p_j)$.

The effect of non-trading on R^2 follows from (i) the effect of non-trading on the estimate of β and (ii) its effect on the variance of observed returns relative to the variance of true returns. For true returns, R^2 equals

$$R_j^2 = \frac{Var\left[R_{jt}^*\right] - Var\left[\varepsilon_{jt}\right]}{Var\left[R_{jt}^*\right]} = \frac{\beta_j{}^2 Var[R_{mt}]}{Var\left[R_{jt}^*\right]}. \tag{5a}$$

14. This assumption can be expected to hold as long as the firm analyzed is small relative to the market and market returns are not auto-correlated through time.

For observed returns, R^2 equals

$$R_j^{2,observed} = \frac{Var[R_{jt}] - Var[\varepsilon_{jt}]}{Var[R_{jt}]} = \frac{b_j^2 Var[R_{mt}]}{Var[R_{jt}]} = \frac{q_j^2 \beta_j^2 Var[R_{mt}]}{Var[R_{jt}]}. \tag{5b}$$

A comparison of these equations suggests that:

if $Var[R_{jt}] = Var[R_{jt}^*]$: $R_j^{2,observed} < R_j^2$ when $q_j < 1$.

if $Var[R_{jt}] > Var[R_{jt}^*]$: $R_j^{2,observed} < R_j^2$.

if $Var[R_{jt}] < Var[R_{jt}^*]$: $R_j^{2,observed} > R_j^2$ when $Var[R_{jt}] < q_j^2 Var[R_{jt}^*]$ or $\frac{Var[R_{jt}]}{Var[R_{jt}^*]} < q_j^2$.

These comparisons suggest that whether zero returns increase or decrease R^2 depends on the ratio of the variance of observed returns ($Var[R]$) to the variance of true returns ($Var[R^*]$). Scholes and Williams (1977) and Lo and MacKinlay (1990) derive that non-trading increases the variance of observed returns, which implies that observed R^2s are biased downward in the presence of zero returns: $R_j^{2,observed} < R_j^2$.[15]

In Appendix 1, we present simulations to further illustrate how non-trading biases the measurement of R^2 downward, and to quantify the strength of the measurement bias in typical research applications. In these simulations, we randomly replace a fraction of nonzero-return days with zero returns and correct the first subsequent trading (nonzero-return) day for the previous non-trading according to equation (2). The simulations confirm that measurement bias leads to strong correlations between non-trading frequencies and R^2 measurements.

Common adjustments for non-trading biases

Previous studies account for non-trading effects by, for example, including a lag of market returns in the market model (e.g., Piotroski and Roulstone 2004), by including multiple lag and lead terms (e.g., Jin and Myers 2006), or by focusing on returns measured in lower frequency (e.g., weekly instead of daily returns).[16] However, these methods were originally designed to obtain consistent estimates of beta by addressing the errors-in-variables problem in the market model (e.g., Scholes and Williams 1977; Dimson 1979), not to correct R^2 measurement. Therefore, the extent to which these methods can address non-trading biases in R^2 measurements is unclear.

Consider a stylized example where the true R^2 equals 0.50 and the true β equals 1.0. A probability of non-trading of $p = 0.2$ leads to an expected beta estimate of 0.8 instead of 1.0 in a simple market model regression. Assuming the variance of observed returns equals that of true returns ($Var[R_{jt}] = Var[R_{jt}^*]$), expected R^2 equals 0.32 ($= 0.5 \times 0.8^2$). Assuming that non-trading days are not consecutive, adding one lag of market returns to the market model would correct the

15. The formulas that explain the variance of observed returns relative to the variance of true returns, derived in Scholes and Williams (1977) and Lo and MacKinlay (1990), imply that the upward bias in observed return variance is small. For example, based on Lo and MacKinlay's (1990) equation (2.11), we can derive that with $p = 0.5$, an expected daily return of 0.0005, and an SD of daily returns of 0.025 (Campbell et al. 1997, 90), the expected ratio of the observed to true return variance equals 1.0008. Accordingly, Scholes and Williams (1977, 314) conclude that with daily returns "measured variances . . . closely approximate true variances."

16. For instance, Durnev et al. (2003, 800) argue that "[w]e use weekly returns because CRSP daily returns data report a zero return when a stock is not traded on a given day. Although some small stocks may not trade for a day or more, they generally trade at least once every few days. Weekly returns are therefore less likely to be affected by such thin trading problems." Morck et al. (2000) use biweekly returns. Bartram et al. (2012) exclude from their analyses firm-years with more than 30 percent of zero weekly returns in the prior year. Still, the use of daily returns without lag and lead terms is common (e.g., Crawford et al. 2012; Peterson et al. 2015; Israeli et al. 2017).

measurement of β because the expected coefficient on lagged market returns would equal 0.2, and the expected sum of the coefficients would be a consistent estimate of the true β of 1.0 (= 0.8 + 0.2). Further assuming stationarity and no serial dependence in market returns, however, the expected $R^{2,observed}$ would still be lower than the true R^2 because

$$R_j^{2,observed} = \frac{Var[b_1R_{mt} + b_2R_{mt-1}]}{Var[R_{jt}]} = \frac{b_1^2\,Var[R_{mt}] + b_2^2\,Var[R_{mt-1}] + 2b_1b_2Cov[R_{mt}, R_{mt-1}]}{Var[R_{jt}]}$$
$$= \frac{b_1^2\,Var[R_{mt}] + b_2^2\,Var[R_{mt-1}]}{Var[R_{jt}]} = \frac{0.8^2 Var[R_{mt}] + 0.2^2 Var[R_{mt-1}]}{Var[R_{jt}]} = \frac{0.68\,Var[R_{mt}]}{Var[R_{jt}]}.$$

That is, given that true $R_j^2 = Var[R_{mt}]/Var\left[R_{jt}^*\right]$, $Cov[R_{mt}, R_{mt-1}] = 0$, and $Var\left[R_{jt}\right] = Var\left[R_{jt}^*\right]$, the measurement of R^2 is biased downward by 32 percent even if the addition of a lag term leads to consistent estimates of beta.[17]

A similar argument applies to the use of return data of lower frequency. For example, while the use of weekly returns addresses the biases when non-trading occurs only in the middle of the week, R^2 measurement can still be biased when non-trading leads a firm's true daily return to be incorporated into observed returns in the following week. The inclusion of lagged weekly market returns allows the researcher to obtain a consistent estimate of beta, but this does not effectively address the reduction in R^2 caused by the deviation between true and observed firm-level returns. Our simulations in Appendix 1 confirm that commonly applied remedies are not effective. Specifically, although we find that biases become smaller with common remedies, the correlations between induced non-trading and R^2 bias remain nontrivial.[18]

3. Data and R^2 measurement

Sample selection details

To test the implications of the arguments laid out in the previous sections, we obtain a global set of firm-year observations over the period 1990–2012 from the 50 countries examined in Bartram et al. (2012). To be included in the sample, a firm's primary listing should be in the same country as that in which the firm is domiciled. We exclude all secondary listings and focus on firms listed on a country's primary exchange.[19] For U.S. firms, market data, including daily returns, and

17. Relaxing the assumption of no autocorrelation in the market returns, we would only obtain the true R^2 with perfect correlation between current and lagged market returns, which is unrealistic. This follows because $2b_1b_2Cov[R_{mt}, R_{mt-1}] = 2b_1b_2\rho[R_{mt}, R_{mt-1}]\sigma[R_{mt}]\sigma[R_{mt-1}]$. If $\rho[R_{mt}, R_{mt-1}] = 1$ and given $\sigma[R_{mt}] = \sigma[R_{mt-1}]$, this equals $2b_1b_2Var[R_{mt}]$ and leads to $b_1^2Var[R_{mt}] + b_2^2Var[R_{mt-1}] + 2b_1b_2Var[R_{mt}] = Var[R_{mt}]$ when $b_1 + b_2 = 1$. Relaxing the assumption of equal variances of true and observed returns, the prediction from Scholes and Williams (1977) and Lo and MacKinlay (1990) that $Var\left[R_{jt}\right] > Var[R_{jt}^*]$ would further reduce R^2 relative to its true value, although this effect would be relatively small.

18. Another approach to addressing non-trading problems is to use "trade-to-trade" methods (e.g., Dimson and Marsh 1983) by aggregating zero-return days with their consecutive trading days and estimating market models based on observations of different lengths. Because of the way we constructed our simulations, this approach fixes the measurement of R^2. In actual data, however, this approach does not work effectively because the exact time of trading is unknown (Dimson 1979; Cohen et al. 1983), and a stock can have nonzero returns while still being relatively illiquid. For example, if illiquidity allows only half of the firm's true returns to be incorporated into price on the same day, we would observe nonzero returns but R^2 would still be biased. This point also highlights the limitations of our characterization of the measurement problems discussed in this section because of our focus on the extreme case in which illiquidity leads to non-trading and a zero return.

19. By primary exchange, we mean the country's exchange with the largest number of listed companies (Watanabe et al. 2013). We allow for multiple primary exchanges in some countries (Canada, China, India, Japan, Korea, Russia, Taiwan, and the United States) to ensure the vast majority of listed companies from these countries are included in our sample.

accounting data are from CRSP and COMPUSTAT Fundamentals Annual, respectively. For non-U.S. firms, market and accounting data are from COMPUSTAT Global Security Daily and COMPUSTAT Global Fundamentals Annual, respectively (e.g., Li et al. 2014), and we compute daily returns based on daily closing prices in local currency, adjusted for stock splits, dividends, and currency changes (e.g., the switch from domestic currency to Euros for EU firms). Following prior research using international security returns (e.g., Griffin et al. 2010), we set to missing all daily returns (t and $t - 1$) that are greater than 200 percent and if $R_{t-1} > 1$ and $(1 + R_{t-1})(1 + R_t) - 1 < 0.2$. Data on analyst coverage are from I/B/E/S for both U.S. and non-U.S. firms.

Our initial sample selection criteria lead to a sample of 535,726 firm-year observations across the 50 countries for which we can measure R^2 based on the market model. Requiring additional data for the measurement of the control variables included in our multiple regression framework leads to a final sample of 377,598 firm-year observations.

Variable measurement

For each firm j and calendar year τ, we estimate stock price synchronicity based on the R^2 from an OLS regression of firms' weekly returns, calculated from Wednesday-to-Wednesday closing prices, on contemporaneous (week t) and lagged (week $t - 1$) market returns:

$$R_{jt} = \beta_{0j} + \beta_{1j}R_{mt} + \beta_{2j}R_{mt-1} + \varepsilon_{jt}. \tag{6}$$

Following Morck et al. (2000) and Jin and Myers (2006), we require at least 30 weekly return observations in a firm-year. The return on the market index (R_m) is computed as the value-weighted weekly market return after excluding the weekly return for the firm of interest (e.g., Durnev et al. 2003) to prevent spurious correlations in countries with a small number of firms. Also, we cap the maximum weight for each firm in the market index at 5 percent. For country-years with fewer than 100 firms, the maximum weight is set to 10 percent. Based on the OLS estimation of equation (6), we define R^2 for firm j and year τ as follows:

$$R_{j\tau}^2 = \frac{\hat{\beta}_{1j}^2 Var[R_{mt}] + \hat{\beta}_{2j}^2 Var[R_{mt-1}]}{Var\left[R_{jt}\right]}. \tag{7}$$

Section 2 exploited the fact that zero daily returns occur frequently in international financial markets to demonstrate the effects of illiquidity on R^2. Stock illiquidity, however, is a more complex and continuous construct than is captured, say, by the simple occurrence of non-trading and zero returns. Nevertheless, prior research suggests that measures derived from zero-return frequencies are useful proxies for the more continuous construct of illiquidity, especially in international research settings (Lesmond et al. 1999; Lesmond 2005; Bekaert et al. 2007). Similar to recent studies that examine corporate transparency in international markets (e.g., Daske et al. 2008; Lang et al. 2012), we therefore use the frequency of zero returns as our measure of illiquidity in the main analyses. For a subsample of U.S. data, we alternatively employ the Amihud (2002) measure in section 5.

Does firm-specific information manifest in lower R^2?

A fundamental question in the synchronicity literature is whether low R^2, *on average*, captures firm-specific information impounded in prices or just noise. A necessary condition for low R^2 to capture firm-specific information in prices is that, all else equal, firm-specific information flows should lead to observable drops in R^2. Therefore, we first evaluate whether R^2 is actually able to pick up firm-specific information being impounded in price in our data. Following Roll (1988), we examine R^2 in news versus non-news periods, using annual earnings announcements for the firms in our sample as firm-specific news events. We use weeks with earnings announcements as the event period and the week before the announcement as the non-event period, and require each firm to have at least 10 annual earnings announcements in I/B/E/S during 1990–2012 (n = 8,213).

For each firm we estimate an event- and non-event adjusted R^2 with returns of the event and pre-event week, respectively, using equation (6).[20] Results in the first row of Table 1 ("All firms") indicate that earnings announcements lead to significant reductions in R^2 (t-statistic: −13.05). While statistically significant, however, the difference in average R^2s is modest (average event-R^2s of 0.155 versus non-event-R^2s of 0.200). Albeit slightly stronger, this modest difference is in line with Roll (1988, 564), who finds R^2 from a multifactor (CAPM) model increases only from 0.205 to 0.225 (0.163 to 0.179) when firm-specific news events are ignored. Similarly, the fraction of firms for which event-week R^2 values are lower than non-event week R^2s is only 56.6 percent. To put these numbers in perspective, note that the mean ratio of event- to non-event return variance is 2.321, and 69.6 percent of firms experience an increase in return variance in event weeks.[21] Thus, although these statistics suggest the earnings announcements are associated with nontrivial firm-specific information, the effects revealed by R^2 are not as strong.

To better understand what drives this result, we sort firms into zero-return deciles. The average ratio of event- to non-event return variance exceeds two for each decile, and there is no significant difference between the highest and lowest deciles. Also, untabulated results suggest that zero-return frequencies decrease significantly across all deciles in event weeks, with the strongest absolute decrease for the most illiquid firms. Because increases in liquidity reduce the downward bias in R^2, any *decrease* in R^2 that results from firm-specific information will therefore be masked by the *increase* in R^2 that results from reduced bias. The remaining rows of Table 1 demonstrate this issue. For liquid firms, there are stronger reductions in R^2 associated with earnings announcements (e.g., average of 0.171 versus 0.265 for firms in decile 1, with 62.7 percent of firms experiencing a decline in R^2).[22] For illiquid firms (deciles 9 and 10), however, for which the reduction in downward R^2 bias is largest, the differences in R^2 are insignificant. This result suggests that the R^2 metric does not pick up firm-specific information for all firms.

Returning to the question of whether low R^2 captures information or noise, the results suggest that measurement bias can make low R^2 appear to reflect more noise in prices—and therefore a weaker information environment (Kelly 2014; Li et al. 2014)—when not controlling for the effect of illiquidity. Importantly, firm-specific information does manifest in lower R^2, but stock illiquidity can mask such effects and reduce the R^2 metric's ability to measure firm-specific information in prices. In the next sections, we empirically investigate the strength and consequences of the illiquidity-R^2 relation in a large international sample and assess methods to control for illiquidity in the relation between R^2 and a variable of interest.

4. Large-sample evidence on measurement bias in R^2

Country-level descriptive evidence

Panel A of Table 2 presents descriptive statistics for our sample by country. The sample of 535,726 firm-years with sufficient data to estimate the market model contains 393,164 (142,562) non-U.S. (U.S.) firm-years, and the non-U.S. (U.S.) sample consists of 38,647 (16,232) unique

20. Because the number of weekly observations used in the R^2 calculations can vary by firm, we use adjusted R^2s in these tests, similar to Roll (1988).

21. This ratio of variances is comparable to the abnormal variance ratios used to capture the information content of earnings announcements in the literature; see, for example, Beaver (1968) and Beaver et al. (2018). The key difference is that we rely on total return variance instead of idiosyncratic return variance to ensure we have a measure that does not rely on market model estimations. Untabulated results using idiosyncratic return variance are even stronger, with an average ratio of 2.559 and 71.6 percent of firms experiencing an increase in idiosyncratic return variance.

22. Consistent with this result for liquid firms, Boudoukh et al. (2019) also find strong reductions in R^2 associated with firm-specific news for their sample of large and highly liquid firms in the S&P 500.

TABLE 1
Does firm-specific information manifest in lower R^2?

	$Var_{EA}/Var_{pre\text{-}EA}$	R^2 EA	R^2 Pre-EA	Diff.	t-stat.	% firms lower R^2
All firms	2.321	0.155	0.200	−0.046	−13.05***	56.6
Decile 1 (liquid)	2.803	0.171	0.265	−0.093	−7.74***	62.7
Decile 2	2.278	0.164	0.221	−0.056	−5.20***	56.4
Decile 3	2.257	0.164	0.233	−0.069	−6.57***	59.3
Decile 4	2.104	0.167	0.225	−0.058	−5.25***	59.3
Decile 5	2.286	0.184	0.217	−0.033	−2.97***	55.6
Decile 6	2.160	0.171	0.202	−0.032	−2.86***	54.3
Decile 7	2.120	0.162	0.205	−0.043	−3.81***	56.7
Decile 8	2.228	0.134	0.176	−0.043	−3.85***	57.2
Decile 9	2.159	0.137	0.155	−0.018	−1.60	51.4
Decile 10 (illiquid)	2.818	0.092	0.106	−0.014	−1.31	53.2
Decile 10 − decile 1	0.015	−0.080	−0.159	0.079		−9.4
t-stat.	0.10	−6.53***	−12.99***	4.88***		−3.88***

Notes: This table presents differences in mean R^2s between periods with and without significant firm-specific information events. The firm-specific information events are annual earnings announcements (EA), for which dates are from I/B/E/S for our international sample of firms (see Table 2) with at least 10 annual earnings announcements (n = 8,213) during 1990–2012. Column "Var_{EA} / $Var_{pre\text{-}EA}$" contains the mean ratio of return variance in earnings announcement weeks to return variance in pre-earnings announcement weeks as a measure of information content. The expected value of this ratio is 1 in the absence of new information in earnings announcement weeks. We estimate adjusted R^2 based on the market model in equation (6) by firm separately for the time series of event (EA) weeks and non-event weeks, where a non-event week is the week immediately before each earnings announcement week. Because the number of weekly observations used in the R^2 calculations varies by firm, we use adjusted R^2s in these tests, similar to Roll (1988). Illiquidity decile portfolios are formed by sorting firms based on their overall frequency of zero daily returns. The column with header "% firms lower R^2" indicates the fraction of firms for which R^2 is lower for the event week estimation compared to the non-event week estimation. t-statistics are based on robust standard errors adjusted for heteroskedasticity. *** indicates statistical significance at the level of 0.01.

firms. Because of variation in data coverage in the earlier years of our sample, not all countries have data starting in 1990. We display the sample start dates in the first column of Table 2.

Panel A highlights the substantial variation in R^2 across countries. Average R^2 is lowest for Canada (0.0993) and highest for China (0.3989), in line with prior studies (e.g., Jin and Myers 2006). Assuming relatively high (low) corporate transparency in Canada (China), these differences are consistent with low (high) R^2 reflecting a more (less) informative stock price. However, the average zero-return frequency is also substantially higher in Canada (30.27 percent) than in China (10.33 percent). Similarly, several countries with low R^2 exhibit high levels of non-trading in the data (e.g., Australia, Czech Republic, Ireland, Luxembourg, New Zealand, Peru, the United Kingdom, Venezuela), while others with high R^2 exhibit low levels of non-trading (Egypt, India, Italy, Korea, Taiwan).

The Spearman rank correlation between country-level average zero-return frequency and R^2 equals −0.6703 (untabulated). Panel A of Figure 1 visualizes the strong negative relation across countries, and the trend line suggests this relation may not be linear. The last column of panel A of Table 2 displays the correlation between illiquidity and R^2 *within* countries. Consistent with the cross-country results, the negative correlation between zero-return frequency and R^2 is strong in each of the 50 countries (average of the 50 correlation coefficients: −0.5499). The bottom row displays weighted-averages for the total sample and suggests zero returns are not just a problem in smaller countries. The overall zero-return frequency is around 26 percent for the total sample.[23]

Panel B of Table 2 provides descriptives for the largest countries, as well as for the remaining countries by region, by year, and suggest that variation in R^2 also relates to variation in illiquidity over time. Most strikingly, we find that R^2 is relatively low for the average U.S. firm in the 1990s. While consistent with prior literature, which interprets the low R^2 in the United States as being driven by firm-specific information flow, we also observe a relatively high frequency of non-trading in the earlier years of the U.S. sample.[24] For instance, the low R^2 of 0.1449 in 1990 coincides with a frequency of zero returns of 0.3797. Recent years display a sharp increase in R^2 for the United States. For the last five sample years (2008–2012), the (untabulated) average R^2 for the United States equals 0.2833. This value is higher than the average R^2 observed for each of the other regions in the same period.[25]

The remaining columns of panel B suggest that the other regions display more stable levels of R^2 and illiquidity over time. Interestingly, the low R^2s for Canada and the United Kingdom coincide with high levels of illiquidity. Given that the available data for these countries are characterized by a relatively large fraction of small and infrequently traded firms compared to economies with a small number of large listed firms, this suggests that low country-level synchronicity is at least partly explained by the characteristics of the *average* listed firm in a country.

Firm-year level descriptive analyses

At the firm-year level, we test the relation between synchronicity and illiquidity in a multiple regression framework. Because R^2 is bounded between 0 and 1, we transform R^2 to an unbounded continuous variable *Synch* (e.g., Morck et al. 2000):

$$Synch_{j\tau} = \ln\left(\frac{R_{j\tau}{}^2}{1-R_{j\tau}{}^2}\right), \tag{8}$$

where $j\tau$ refers to firm-year.

We control for several previously identified determinants of synchronicity (Chan and Hameed 2006; Ferreira and Laux 2007; Fernandes and Ferreira 2008). Specifically, we control for profitability (*ROE*), volatility of profitability (*VROE*), leverage (*Leverage*), market-to-book (*MTB*), market capitalizations in US$ (*Size*), dividends (*DD*), and trading volume (*Volume*).[26] We also include

23. In contrast to R^2, the correlation between zero-return frequency and idiosyncratic return volatility (σ_ε^2) is much weaker. The cross-country Spearman correlation equals 0.0574, while the average within-country correlation equals 0.0676. These weaker correlations underscore the differential effect of illiquidity on R^2 versus idiosyncratic volatility measurement and that R^2 versus idiosyncratic volatility are not interchangeable (Li et al. 2014). This highlights the important difference between our work and Han and Lesmond (2011).

24. Note that for the U.S. sample the earlier years in our sample are unlikely driven by a sample selection bias, given that our data sources (CRSP and COMPUSTAT) started coverage of U.S. firms long before the 1990s.

25. If high R^2 indicates low price informativeness due to a weak information environment, these data would suggest that the United States has changed to a relatively weak information environment in recent years. This interpretation appears to be in stark contrast with prior evidence indicating that regulatory reforms, such as the Sarbanes-Oxley Act of 2002 (SOX), have succeeded in increasing corporate transparency. For example, prior research suggests that firms' financial information became more transparent after SOX as firms engaged in less earnings management and disclosed material weaknesses in internal control over financial reporting, reducing information asymmetries between managers and outside investors (e.g., Cohen et al. 2008; Skaife et al. 2013, respectively). To the extent that increased transparency and reduced private information for managers decreases synchronicity (Jin and Myers 2006), the pattern of a strong increase in synchronicity in the United States post-SOX would imply lower levels of transparency and increased private information for managers.

26. Note that several of the determinants are based on accounting measures that are a function of the accounting standards applied in preparing a firm's financial statements. When sample firms follow different sets of accounting standards, then a determinant based on accounting measures may not be comparable across firms (Gordon et al. 2013). We include these determinants in our models to align with prior synchronicity research, and use country fixed effects to control for systematic differences in these variables that arise from differences in accounting standards across countries.

TABLE 2
Country-level descriptive statistics

Panel A: Synchronicity and zero-return frequency by country

Country	Sample start	Firm-years	Median number firms per year	Mean R^2	Mean ZR	Rank corr. R^2 and ZR
Argentina	1995	1,119	64	0.2432	0.4283	−0.6953
Australia	1990	21,505	1,549	0.1026	0.4320	−0.4963
Austria	1990	1,963	89	0.1527	0.3791	−0.6130
Belgium	1990	2,524	129	0.1798	0.2975	−0.5723
Brazil	1991	4,141	256	0.1484	0.3840	−0.5941
Canada	1990	49,031	2,998	0.0993	0.3027	−0.3920
Chile	1992	2,603	151	0.1456	0.6097	−0.7013
China	1994	15,361	1,572	0.3989	0.1033	−0.3343
Colombia	1997	471	29	0.2189	0.5032	−0.7313
Czech Republic	1996	432	31	0.1172	0.4927	−0.6018
Denmark	1990	3,211	179	0.1549	0.3913	−0.6243
Egypt	1998	626	51	0.2930	0.2101	−0.2389
Finland	1990	2,243	126	0.2040	0.2927	−0.5991
France	1990	12,784	677	0.1453	0.2820	−0.6402
Germany	1990	12,472	670	0.1287	0.2746	−0.5977
Greece	1995	3,355	243	0.2850	0.2060	−0.6163
Hong Kong	1990	5,470	251	0.2289	0.3121	−0.6186
Hungary	1996	381	23	0.2392	0.1850	−0.5884
India	1990	18,325	1,436	0.2500	0.1277	−0.3825
Indonesia	1991	5,779	309	0.1445	0.5887	−0.6044
Ireland	1990	1,140	51	0.1435	0.4453	−0.5284
Israel	1995	3,176	221	0.2413	0.2797	−0.5789
Italy	1990	5,244	257	0.2526	0.1155	−0.4324
Japan	1990	67,214	3,369	0.2327	0.2098	−0.4329
Korea	1990	19,872	1,433	0.2163	0.1052	−0.2211
Luxembourg	1996	274	18	0.1273	0.4348	−0.4679
Malaysia	1990	14,897	915	0.2465	0.3501	−0.6195

(The table is continued on the next page.)

TABLE 2 (continued)

Panel A: Synchronicity and zero-return frequency by country

Country	Sample start	Firm-years	Median number firms per year	Mean R^2	Mean *ZR*	Rank corr. R^2 and *ZR*
Mexico	1991	1,899	98	0.1911	0.4488	−0.7311
Morocco	1999	655	69	0.1881	0.4139	−0.5745
Netherlands	1990	3,570	160	0.1948	0.2389	−0.6436
New Zealand	1990	1,709	103	0.1049	0.4681	−0.5131
Norway	1990	2,904	162	0.1731	0.3772	−0.6334
Pakistan	1995	3,364	226	0.1965	0.3536	−0.6229
Peru	1997	874	64	0.1398	0.6653	−0.6446
Philippines	1993	3,668	197	0.1473	0.5903	−0.5971
Poland	1996	3,419	304	0.1974	0.2091	−0.4655
Portugal	1994	1,030	53	0.1741	0.3481	−0.6710
Russian Federation	1997	1,537	157	0.2015	0.4194	−0.5673
Singapore	1990	8,466	564	0.2013	0.4575	−0.6459
South Africa	1990	5,814	292	0.1266	0.4378	−0.5662
Spain	1990	2,983	135	0.2343	0.2318	−0.6083
Sri Lanka	1998	2,643	176	0.2313	0.5448	−0.6129
Sweden	1990	5,224	291	0.1953	0.2336	−0.5162
Switzerland	1990	3,887	229	0.1888	0.2708	−0.6378
Taiwan	1992	15,912	1,269	0.2896	0.1497	−0.3325
Thailand	1990	8,747	396	0.1986	0.3932	−0.6183
Turkey	1995	3,178	202	0.3564	0.1793	−0.1884
United Kingdom	1990	35,650	1,617	0.1445	0.4955	−0.5274
United States	1990	142,562	6,181	0.1552	0.1551	−0.5321
Venezuela	1997	418	26	0.1246	0.7399	−0.5237
Total	1990	535,726	25,875	0.1834	0.2594	−0.4893

(The table is continued on the next page.)

TABLE 2 (continued)

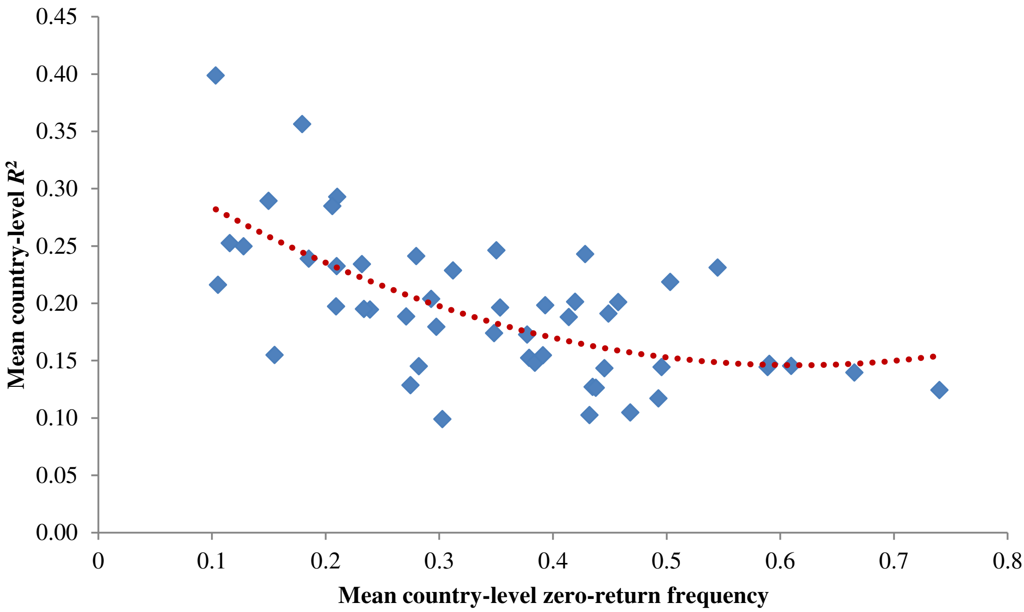
Panel B: Synchronicity and zero-return frequency by sample year and region

	U.S. Mean		Canada Mean		UK Mean		Europe^{cont} Mean		Japan Mean		Asia^{other} Mean		Australasia Mean		Africa Mean		Latin America Mean	
Year	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*	R^2	*ZR*
1990	0.14	0.38	0.10	0.36	0.22	0.53	0.32	0.31	0.38	0.18	0.41	0.38	0.11	0.61	0.18	0.59	–	–
1991	0.14	0.37	0.12	0.37	0.20	0.57	0.25	0.32	0.29	0.23	0.29	0.44	0.12	0.55	0.15	0.63	0.27	0.18
1992	0.10	0.31	0.08	0.36	0.25	0.58	0.21	0.35	0.42	0.27	0.28	0.38	0.12	0.53	0.18	0.66	0.25	0.38
1993	0.08	0.27	0.06	0.31	0.11	0.57	0.18	0.30	0.29	0.22	0.24	0.34	0.11	0.44	0.14	0.63	0.17	0.58
1994	0.10	0.27	0.10	0.31	0.17	0.56	0.21	0.33	0.24	0.28	0.29	0.32	0.14	0.43	0.15	0.61	0.30	0.43
1995	0.08	0.26	0.07	0.31	0.11	0.62	0.15	0.34	0.27	0.28	0.30	0.35	0.10	0.46	0.13	0.64	0.32	0.37
1996	0.10	0.24	0.06	0.25	0.10	0.58	0.13	0.32	0.20	0.27	0.23	0.29	0.12	0.40	0.11	0.49	0.18	0.39
1997	0.10	0.20	0.09	0.22	0.10	0.53	0.17	0.26	0.18	0.30	0.31	0.28	0.17	0.38	0.16	0.36	0.24	0.45
1998	0.17	0.16	0.10	0.30	0.16	0.49	0.23	0.26	0.22	0.29	0.30	0.32	0.11	0.42	0.18	0.31	0.24	0.52
1999	0.07	0.16	0.05	0.31	0.09	0.49	0.13	0.29	0.12	0.23	0.25	0.29	0.07	0.38	0.12	0.34	0.20	0.51
2000	0.11	0.15	0.06	0.26	0.10	0.46	0.16	0.25	0.11	0.23	0.22	0.31	0.08	0.33	0.15	0.37	0.14	0.55
2001	0.14	0.10	0.07	0.32	0.18	0.49	0.20	0.27	0.19	0.22	0.27	0.31	0.09	0.36	0.13	0.43	0.15	0.60
2002	0.19	0.07	0.08	0.33	0.16	0.52	0.20	0.31	0.21	0.21	0.24	0.29	0.08	0.42	0.08	0.47	0.14	0.61
2003	0.16	0.06	0.06	0.29	0.14	0.51	0.18	0.30	0.16	0.19	0.23	0.27	0.07	0.46	0.11	0.48	0.12	0.57
2004	0.17	0.05	0.07	0.25	0.10	0.49	0.14	0.26	0.22	0.15	0.21	0.26	0.06	0.42	0.09	0.44	0.13	0.54
2005	0.14	0.05	0.07	0.23	0.12	0.47	0.13	0.23	0.13	0.13	0.21	0.24	0.08	0.41	0.14	0.38	0.12	0.54
2006	0.15	0.04	0.11	0.20	0.15	0.45	0.20	0.20	0.29	0.13	0.20	0.24	0.09	0.39	0.20	0.32	0.15	0.50
2007	0.23	0.04	0.12	0.20	0.16	0.44	0.20	0.19	0.22	0.14	0.21	0.18	0.11	0.35	0.16	0.29	0.15	0.40
2008	0.28	0.04	0.17	0.36	0.16	0.46	0.27	0.24	0.35	0.16	0.37	0.22	0.15	0.44	0.18	0.33	0.20	0.42
2009	0.28	0.04	0.12	0.41	0.13	0.46	0.18	0.28	0.17	0.22	0.25	0.23	0.10	0.49	0.17	0.40	0.13	0.46
2010	0.29	0.04	0.11	0.36	0.14	0.45	0.18	0.26	0.25	0.23	0.22	0.22	0.11	0.46	0.14	0.39	0.15	0.44
2011	0.36	0.03	0.17	0.33	0.22	0.41	0.27	0.27	0.40	0.22	0.29	0.23	0.14	0.46	0.20	0.41	0.20	0.45
2012	0.21	0.04	0.12	0.40	0.13	0.42	0.14	0.30	0.23	0.21	0.21	0.24	0.09	0.52	0.14	0.40	0.12	0.49
Total	0.16	0.16	0.10	0.30	0.14	0.50	0.19	0.27	0.23	0.21	0.25	0.25	0.10	0.43	0.15	0.42	0.17	0.49

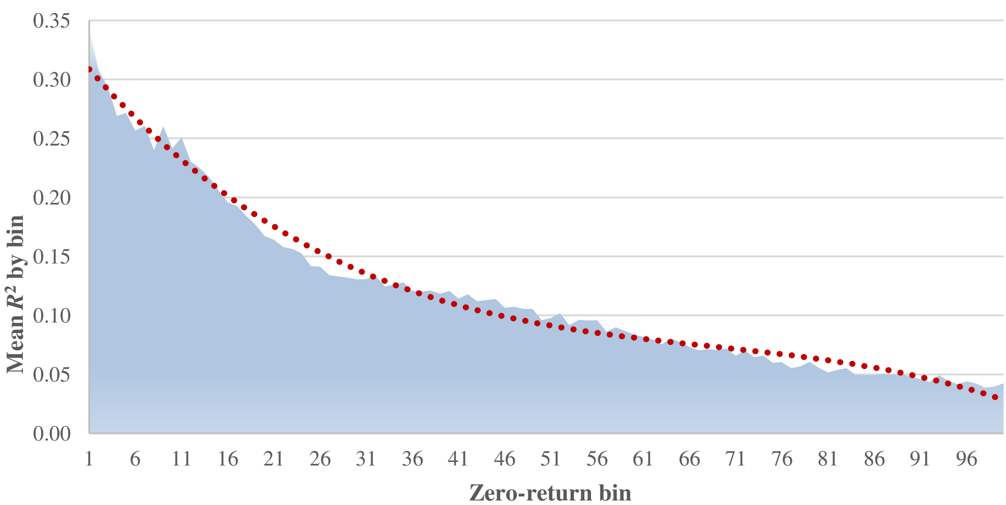
Notes: This table presents country-level descriptive statistics for our global sample of 535,726 firm-year observations for the period 1990–2012. For the analyses in subsequent tables, we rely on a subset of 377,598 firm-years with available data on relevant control variables. Data for non-U.S. firms are obtained from COMPUSTAT Global Fundamentals Annual and Security Daily for accounting data and stock price data, respectively. Data for U.S. firms are obtained from CRSP and COMPUSTAT North America. For each country, we include only those companies that are domiciled and listed in that country. R^2 is the explanatory power obtained from a firm-year specific market model regression using weekly returns, including contemporaneous and lagged market returns that are value-weighted market returns excluding the firm of interest (see equation (6)). Weekly returns are computed from Wednesday to Wednesday. *ZR* is the fraction of trading days for which a zero return is observed for a specific company-year.

Figure 1 Relation between stock price synchronicity and zero-return frequencies

Panel A: Average country-level stock price synchronicity and zero-return frequencies



Panel B: Relation between firm-year stock price synchronicity and zero-return frequencies



Notes: Panel A presents the relation between country-level average R^2 (estimated based on the market model in equation (6)) and country-level average zero-return frequency for the 50 countries listed in Table 2. The horizontal axis displays the mean country-level zero daily return frequency, whereas the vertical axis displays the mean country-level R^2. Panel B presents the average R^2 for each bin of with 0.01 in [0.00, 1.00) of zero-return frequency in our global sample of 377,598 firm-years with sufficient data (see Table 3). In panel A (panel B), the trend lines represent a fitted equation estimated including two (three) statistically significant (p-value < 0.05) polynomial terms.

country-year and industry fixed effects in our analyses. Inclusion of the country-year fixed effects picks up the substantial variation in R^2, as well as the differences in temporal variation in R^2, across countries. We use *ZeroReturn*, defined as the fraction of zero-return days in a firm-year, as our main variable of interest.

Panel A of Table 3 presents descriptive statistics for the sample of 377,598 firm-year observations having data available for all variables. We winsorize continuous variables at the 1st and 99th percentiles to reduce the influence of extreme observations. Because of the skewness in these variables, we also take the natural logarithm of variables *VROE*, *MTB*, *Size*, and *Volume* in all subsequent analyses.

Panel B presents descriptive statistics (means) for stock price synchronicity by 10 bins of zero-return frequency. As the frequencies per row reveal, the distribution of zero-return rates is right-skewed with 38.5 percent of observations falling in the [0.00–0.10) bin. Next, the descriptives reveal substantial variation in estimated market model coefficients. While illiquid firms are arguably riskier (e.g., Amihud and Mendelson 1986), estimates of β_1 are monotonically decreasing with zero-return frequencies. This pattern is remarkably similar to that found in the simulations in Appendix 1 and as predicted in section 2. Estimates of β_2 increase with zero-return frequency, which is consistent with the increasing importance of including a lag term, but decrease for zero-return frequencies greater than 60 percent.

Most importantly, we find a strong negative relation between R^2 and illiquidity, as the most liquid (illiquid) firms have an average R^2 of 0.278 (0.043). Consistent with our simulations in Appendix 1, the relation between zero-return frequencies and R^2 is nonlinear. Panel B of Figure 1 provides graphical evidence to support this interpretation. That is, for liquid firms, the slope of the relation between R^2 and zero-return frequency is much steeper than for illiquid firms. We find a similar nonlinearity for the unbounded continuous variable *Synch*. Panel C presents descriptive evidence on the means of our control variables by zero-return bins. We find that the variables linked to *Synch* also vary substantially with zero-return frequency, mostly in a nonlinear fashion.

Panel A of Table 4 reports results of OLS regressions for both the pooled sample of firm-years and for each country separately in a Fama and MacBeth (1973) cross-sectional regression framework. The results suggest that *Synch* is significantly positively associated with *Size* and *Volume*, and significantly negatively associated with *VROE* and *MTB*.[27] The coefficient on *ROE* (*DD*) is significantly positive only in the pooled (by-country) estimation. Out of these variables, *Size* and *Volume* are most strongly related to synchronicity, as indicated by their *t*-statistics and significant coefficients in 36 and 28 country-level estimations, respectively. We also find that the strongest determinant of *Synch* is *ZeroReturn*, as indicated by a significant negative coefficient for 49 of 50 countries.

Overall, the results in Tables 2 through 4 suggest that R^2 is strongly negatively related to illiquidity *across* and *within* countries.[28] These results support the implications derived from our econometric modeling and suggest that illiquidity is empirically strongly correlated with R^2 in common international samples. Moreover, we find that the empirical relation between zero-return frequencies and R^2 is nonlinear. As discussed in subsequent analyses, such nonlinearity can have important implications for attempts at controlling for illiquidity.

5. Controlling for illiquidity

The effects of common adjustments for the non-trading problem in our data

Section 2 discussed conceptually why common remedies to address the implications of non-trading in the market model are unlikely to resolve the measurement problem with R^2. In this section, we *empirically* evaluate how different market model estimations affect the relation between R^2 and

27. Note that *Volume* partly captures the effect of illiquidity on synchronicity measurement. However, because trading volume is a noisy proxy for liquidity and for consistency with prior work (Chan and Hameed 2006), we include trading volume as a determinant to investigate the incremental impact of illiquidity.

28. All results presented in this section are qualitatively similar when we use the extended market model for international samples of Jin and Myers (2006), which includes two leads and lags of market returns.

TABLE 3
Firm-year level descriptive statistics

Panel A: Descriptive statistics for variables used in regressions

Variable	Mean	SD	p5	Q1	Median	Q3	p95
R^2	0.197	0.174	0.009	0.054	0.143	0.302	0.556
Synch	−1.950	1.505	−4.730	−2.873	−1.788	−0.839	0.226
ZeroReturn	0.229	0.232	0.012	0.063	0.142	0.313	0.766
ROE	−0.003	0.367	−0.566	0.001	0.067	0.139	0.292
VROE	0.242	0.628	0.010	0.031	0.069	0.171	0.951
Leverage	0.212	0.189	0.000	0.036	0.179	0.339	0.572
MTB	2.441	3.821	0.276	0.771	1.377	2.524	7.562
SIZE (mln)	1,163	3,653	6	33	124	543	5,448
DD	0.497	0.500	0	0	0	1	1
Volume (mln)	244	777	0	2	17	103	1,278

Panel B: Zero-return frequencies and mean stock price synchronicity by levels of zero-return frequency

Interval	No. obs.	% of total	*ZeroReturn*	β_1	β_2	R^2	*Synch*
[0.00, 0.10)	145,429	38.5	0.048	1.073	0.087	0.278	−1.267
[0.10, 0.20)	83,616	22.1	0.144	0.889	0.153	0.210	−1.773
[0.20, 0.30)	49,311	13.1	0.245	0.736	0.180	0.146	−2.322
[0.30, 0.40)	28,389	7.5	0.345	0.645	0.190	0.125	−2.520
[0.40, 0.50)	18,427	4.9	0.446	0.567	0.198	0.110	−2.671
[0.50, 0.60)	14,473	3.8	0.548	0.493	0.201	0.093	−2.863
[0.60, 0.70)	11,946	3.2	0.649	0.396	0.197	0.075	−3.079
[0.70, 0.80)	10,335	2.7	0.748	0.306	0.170	0.062	−3.294
[0.80, 0.90)	8,727	2.3	0.848	0.205	0.135	0.051	−3.504
[0.90, 1.00)	6,945	1.8	0.947	0.081	0.056	0.043	−3.708
Total	377,598	100.0	0.229	0.828	0.138	0.197	−1.950

Panel C: Means of control variables by levels of zero-return frequency

Interval	*ROE*	ln(*VROE*)	*Leverage*	ln(*MTB*)	ln(*SIZE*)	*DD*	ln(*Volume*)
[0.00, 0.10)	0.039	−2.639	0.217	0.558	17.49	0.546	17.75
[0.10, 0.20)	−0.008	−2.592	0.218	0.306	17.27	0.485	16.72
[0.20, 0.30)	−0.025	−2.519	0.216	0.232	16.02	0.471	15.92
[0.30, 0.40)	−0.039	−2.465	0.211	0.106	16.20	0.467	15.63
[0.40, 0.50)	−0.047	−2.423	0.202	0.047	16.74	0.479	15.40
[0.50, 0.60)	−0.052	−2.378	0.191	0.046	16.96	0.481	15.21
[0.60, 0.70)	−0.058	−2.342	0.188	0.026	17.02	0.470	14.80
[0.70, 0.80)	−0.062	−2.334	0.184	−0.019	16.95	0.443	14.26
[0.80, 0.90)	−0.069	−2.271	0.181	−0.105	16.80	0.382	13.54
[0.90, 1.00)	−0.046	−2.321	0.186	−0.156	16.76	0.280	12.27
Total	−0.003	−2.547	0.212	0.320	17.04	0.497	16.52

Notes: R^2 is based on estimations of equation (6) using weekly returns. *Synch* is the logarithmic transformation of R^2 (see equation (7)). *ZeroReturn* is the fraction of zero-return days in a firm-year. *ROE* is return on equity, defined as the ratio of income before extraordinary items to common equity. *VROE* is the SD of *ROE* over the past five (minimum three) years. *Leverage* is the ratio of long-term debt to total assets for the most recently completed fiscal year before the current calendar year. *MTB* is the market-to-book ratio, computed as the ratio of market capitalization in December of the previous calendar year to common equity for the most recent fiscal year. *Size* is market capitalization in December of the previous calendar year in US$. *DD* is an indicator set equal to one if the firm paid dividends during the most recent fiscal year, and zero otherwise. *Volume* is the sum of the number of shares traded during the year. All continuous variables are winsorized at the 1st and 99th percentiles.

illiquidity. Panel B of Table 4 presents the results. Using a market model with daily returns, the correlation between *ZeroReturn* and *Synch* is −0.591. Consistent with the use of weekly returns reducing the problem, the correlation becomes smaller but remains substantial at −0.465. Adding lags and leads further reduces the correlations, but only marginally. For example, with four lags (e.g., Hou and Moskowitz 2005; Callen et al. 2013), the correlation is −0.429. These analyses confirm our discussion in section 2, as well as the simulations in Appendix 1, and suggest that common adjustments applied to the market model estimations do not eliminate the bias in R^2. Hence, researchers should look for alternative approaches to attenuate the illiquidity bias in R^2.

Linear versus nonlinear controls for illiquidity

A simple approach to attenuate the bias in R^2 is to control for a measure of illiquidity in a regression. A disadvantage is that this approach assumes a linear functional form in the relation between synchronicity and illiquidity, while the descriptives in Table 3, Figure 1, and Appendix 1 suggest this assumption may not hold. One solution to address this nonlinearity is to include polynomials of the illiquidity variable in the regressions. However, the number of polynomials needed using this approach is unclear and can vary by setting. For example, the trend lines displayed in panels A and B of Figure 1 are based on the use of different numbers (two and three, respectively) of statistically significant (p-value < 0.05) polynomial terms.

An alternative solution is to use a flexible control variables approach, such as LSDV regressions where the control variable of interest is split up into n dummy variables that each take on a value of one for specific values of the control variable, and zero otherwise (e.g., Gormley and Matsa 2014).[29] Assuming each firm-year has 250 trading days, an example of this approach would be to include 251 dummy variables that each capture the number of zero-return days in the year. In this case, the zero-return variable would be “fully saturated” and the functional form of the relation between synchronicity and illiquidity becomes fully flexible (e.g., Angrist and Pischke 2009).[30]

To explore the merit of a flexible control variables approach, we return to the regressions in Table 4 and compare the incremental explanatory powers for alternative model specifications incorporating the different control variable approaches. Results in panel C suggest that the inclusion of controls based on *ZeroReturn* adds meaningful explanatory power (adjusted R^2) to both the pooled and by-country estimations. While inclusion of two and five dummy variables (bin widths of 0.50 and 0.20, respectively) for zero returns adds less explanatory power than the linear control variable approach, the inclusion of 10 dummy variables (bin width of 0.10) achieves about the same explanatory power. Moving to more dummy variables with smaller bin widths, we find that the incremental explanatory power increases relative to the linear control variable approach. For example, in the by-country analyses, explanatory power increases by 0.059 for n = 100, and by 0.067 points for n = 250. In contrast, inclusion of a linear control variable increases explanatory power by only 0.044.

These statistics suggest that allowing for a more flexible functional form of the relation between synchronicity and illiquidity allows the regression to pick up incremental variation in synchronicity—variation that might correlate with a variable of interest and result in an omitted variable problem if not controlled for. An important caveat to *any* control variable approach is that we should be careful not to over-control and thereby eliminate variation in R^2 that is relevant to the research question. Such a situation might occur when illiquidity is part of the conceptual mechanism of interest—for example, as a mediating variable. In Appendix 2, we discuss this issue in more detail.

29. Although the terms are technically similar, we mostly refer to LSDV and dummy variables instead of fixed effects in the paper because researchers typically use fixed effects approaches in different ways to control for *unobserved* heterogeneity across a specific dimension, such as when including firm or industry fixed effects.

30. Note that the number of dummy variables included in this hypothetical regression is a *maximum* of 251. Because not all unique values of the zero-return variable may be represented in the data, multiple dummy variables are likely to be dropped from the estimation. Statistical programs such as Stata take care of this automatically.

TABLE 4
Firm-year multiple regressions: Relation between illiquidity and stock price synchronicity

Panel A: Multiple regressions of stock price synchronicity on determinants

Estimation:	Pooled	50 country-level cross-sectional regressions		
Dependent variable:	*Synch*	*Synch*	$<0^{**}$	$>0^{**}$
ZeroReturn	−1.920	−2.317	49	0
	(−21.11)***	(−21.39)***		
ROE	0.058	0.038	0	8
	(3.16)***	(1.59)		
ln(*VROE*)	−0.023	−0.014	11	2
	(−4.03)***	(−2.11)**		
Leverage	0.011	0.013	4	7
	(0.30)	(0.23)		
ln(*MTB*)	−0.080	−0.065	22	1
	(−4.75)***	(−5.21)***		
ln(*SIZE*)	0.176	0.142	0	36
	(14.43)***	(12.99)***		
DD	0.031	0.034	0	8
	(1.33)	(2.46)**		
ln(*Volume*)	0.085	0.062	0	28
	(21.45)***	(7.24)***		
Country-year fixed effects	Yes	No		
Year fixed effects	No	Yes		
Industry fixed effects	Yes	Yes		
n	377,596	377,598		
Adjusted R^2	0.476			
Average adjusted R^2		0.466		

Panel B: Correlations between different synchronicity measurements and illiquidity measures

Return interval:	Daily	Weekly	Weekly	Weekly	Weekly
Leads/lags:	—	—	1 lag	2 lags/leads	4 lags
	Synch	*Synch*	*Synch*	*Synch*	*Synch*
Correlation with *ZeroReturn*	−0.592	−0.465	−0.458	−0.431	−0.429

Panel C: Incremental explanatory power for different methods of controlling for zero-return frequencies

		LSDV with n dummies						
	Linear control	n = 2	n = 5	n = 10	n = 25	n = 50	n = 100	n = 250
World sample	0.031	0.013	0.025	0.029	0.035	0.038	0.038	0.039
World by country	0.044	0.020	0.038	0.047	0.053	0.056	0.059	0.067

Notes: See the notes to Table 3 for variable definitions. For the pooled regressions, t-statistics (reported in parentheses) are based on robust standard errors adjusted for heteroskedasticity and two-way clustering by firm and year. For the country-level regressions, t-statistics are based on the average and SDs of the 50 coefficient estimates. We only include country-years with at least 10 observations. Industry fixed effects are based on 2-digit SIC industry classifications. Regression intercepts are included but not tabulated. Panel B presents correlations with *ZeroReturn* when synchronicity is measured based on different market model estimations (the model with weekly returns and one lag of market returns is the one used throughout the paper). Panel C presents the increase in explanatory power (adjusted R^2) of the regressions in panel A after including controls for zero-return frequency in different ways, relative to the base regression excluding a control for zero-return frequency. The numbers in the "linear control" column refer to the estimates from panel A. n = {2, 5, 10, 25, 50, 100, 250} refers to the number of dummy variables included in the regression to control for illiquidity. For *ZeroReturn*, dummy variables are created based on membership in the n bins in the [0.00, 1.00) interval of zero-return frequency, where each bin has a width of 1.00/n. The numbers in the column $<0^{**}$ ($>0^{**}$) reflect the number of times a specific coefficient is negative (positive) and statistically significant at $p < .05$ out of the 50 country-level regressions. *** and ** indicate statistical significance at the level of 0.01 and 0.05, respectively.

Implications for research: Analyst coverage tests

Analyst coverage and stock price synchronicity

Using a measure of stock price synchronicity, Chan and Hameed (2006) conclude that emerging-market firms' stock prices reflect more market-wide rather than firm-specific information when analysts provide more coverage. Similarly, Piotroski and Roulstone (2004) conclude that U.S. firms' analyst activities lead to higher synchronicity with market and industry returns. At the same time, evidence by Crawford et al. (2012) suggests the link between analyst coverage and synchronicity is more nuanced. They find that analysts' initiation of coverage is associated with increased synchronicity for firms without preexisting coverage, while synchronicity decreases with initiations for the majority of cases in which firms have preexisting coverage (about 94 percent of their sample). They also find that plausibly exogenous drops in coverage are associated with increases in synchronicity. Hence, these results hint that the marginal effect of analyst coverage on synchronicity may be negative, suggesting the net effect of greater analyst coverage is to increase the amount of firm-specific information impounded in prices.

Given these ambiguous insights, we revisit the analyst coverage and synchronicity link in Table 5. Controlling for illiquidity is likely important, as analysts cover firms that are more liquid and generate higher trading commissions (Alford and Berger 1999) and analyst coverage reduces illiquidity (Roulstone 2003; Kelly and Ljungqvist 2012). Empirically, the Spearman correlation (untabulated) between *ZeroReturn* and the average number of forecasts outstanding (*ANALYSTS*) equals −0.4570 for our sample. We find a similarly strong correlation between *ZeroReturn* and the frequency of forecast revisions (*NREV*): −0.4777.[31] Moreover, descriptive statistics in panel A of Table 5 suggest that this relation is nonlinear. Illiquidity therefore relates to both synchronicity and analyst coverage, and failing to adequately control for illiquidity could affect the inferences from tests of the relation between synchronicity and analyst coverage.

Panel B of Table 5 presents tests of the relation between analyst coverage and synchronicity before and after controlling for illiquidity. Without controls for illiquidity, pooled sample results confirm the positive and significant association as documented in Chan and Hameed (2006) for emerging markets and Piotroski and Roulstone (2004) for the United States, for both analyst coverage variables (columns (1) and (4)). Next, we include *ZeroReturn* as a linear control variable in the regressions (columns (2) and (5)). With controls for illiquidity, the significant positive association between synchronicity and analyst coverage is smaller in magnitude, but remains statistically significant across the estimations.

A different picture emerges when we use the LSDV approach (columns (3) and (6)). Using $n = 100$ dummy variables for *ZeroReturn* that represent 100 bins of width 0.01, the significantly positive coefficients on ln(*NREV*) and ln(*ANALYSTS*) become insignificant. Hence, these results suggest that the relation between analyst coverage and synchronicity is more ambiguous than previously concluded in the literature. Panel C presents the same analyses for the by-country estimations. We find similar results. As before, the LSDV approach also picks up substantially more of the variation in synchronicity than the linear control variable.[32]

To provide more insights into the difference between the traditional linear control variable approach and the LSDV approach, Figure 2 presents estimates and confidence intervals for the coefficient on ln(*NREV*) when using different levels of saturation for the *ZeroReturn* variable (i.e., different numbers and widths of bins). As the figure reveals, we would conclude that analyst coverage and synchronicity are significantly positively related (i) when illiquidity is not controlled

31. We follow Piotroski and Roulstone (2004) in using the number of revisions in one-year-ahead earnings forecasts made during the year. We follow Chan and Hameed (2006) in using the average number of forecasts outstanding during the year.
32. Following Piotroski and Roulstone (2004) and Chan and Hameed (2006), we also assess the robustness of these findings to restricting the sample to observations with positive values for the analyst coverage variables (untabulated). Inferences, and particularly inferences across the different control variable approaches, are qualitatively similar.

for, (ii) when a linear control variable for illiquidity is included, and (iii) when 10 or fewer dummy variables based on *ZeroReturn* bins are included. Using n = 25 bins (width 0.04) or more, however, the positive coefficient loses its statistical significance. As can be assessed by the confidence intervals, this result is not driven by reduced precision. Because the LSDV approach allows for a more flexible relation between illiquidity and both synchronicity and analyst coverage, these results highlight the importance of controlling for nonlinear confounding effects.[33]

Alternative control for illiquidity in settings with low zero-return frequencies

Because of the relatively low frequency of zero returns in the U.S. sample with sufficient data (untabulated 0.1258 versus 0.2635 for the rest of the world), a potential drawback of using the zero-return metric is that it may not have sufficient variation among relatively liquid firms. To assess whether the U.S. results are sensitive to this limited variation in the control variable, we alternatively employ the Amihud (2002) measure as a more continuous proxy for illiquidity.

Table 6 presents results for the United States. In panel A, we first find that controlling for illiquidity using the zero-return metric eliminates the significant positive coefficient on ln(*ANALYSTS*), but not on ln(*NREV*), even with the LSDV approach. To assess the robustness of this result, panel B presents results after controlling for illiquidity using the Amihud measure. The results suggest that including the Amihud measure as a linear control variable eliminates the significant positive coefficient on ln(*ANALYSTS*), but not that on ln(*NREV*). However, when including n = 100 dummy variables that capture the percentiles of the distribution of the Amihud measure, the coefficient becomes insignificant. Therefore, when we combine a more continuous control variable for illiquidity with the LSDV approach, we cannot conclude that analyst coverage is positively associated with synchronicity in the United States after controlling for illiquidity.

Overall, these results suggest that the previously documented positive relation between analyst coverage and synchronicity is sensitive to controls for illiquidity. Besides informing the literature on analyst coverage and synchronicity, our analyses highlight the importance of (i) controlling for illiquidity in studies on synchronicity, (ii) allowing for a more flexible (nonlinear) functional form of the relation between illiquidity and synchronicity, and (iii) controlling for alternative measures of illiquidity in settings where the variation in zero-return frequencies is more limited (e.g., the United States in recent years).

Arguments against a "null" result

A potential argument against finding an insignificant relation in the presence of dummy variables is that the estimation lacks power, because the dummy variables reduce the degrees of freedom. However, the elimination of statistical significance here is unlikely driven by power issues, given that (i) the addition of 100 dummy variables to a regression of 377,598 observations has limited effect on the degrees of freedom and (ii) the coefficient itself changes in magnitude (rather than just the confidence interval becoming wider). However, for smaller samples, such as in the individual country-level estimations, researchers should carefully evaluate the effect of the LSDV approach on both the coefficient estimate *and* its associated confidence interval. For samples with relatively few observations, the inclusion of, for example, 100 dummy variables may have a more dramatic impact than we find here.

At the same time, an argument against the result that the significant positive relation disappears after controlling for illiquidity is that illiquidity is part of the conceptual mechanism being investigated. In such a situation, conditioning the analysis on illiquidity may over-control and

33. To highlight that the LSDV approach drives these results, allowing for a more flexible and nonlinear influence of illiquidity, we also test (untabulated) the consequences of sequentially adding more polynomials of *ZeroReturn* as control variables in a linear regression. Similar to the LSDV approach, we find that incrementally adding more polynomials leads the coefficient on the analyst coverage variables to become insignificantly different from zero (while explanatory power, as measured by adjusted R^2, increases).

TABLE 5

Analyst coverage and stock price synchronicity: Controlling for illiquidity

Panel A: Descriptive statistics for analyst coverage variables by levels of zero-return frequency

	Zero-return frequency interval									
	[0.00, 0.10)	[0.10, 0.20)	[0.20, 0.30)	[0.30, 0.40)	[0.40, 0.50)	[0.50, 0.60)	[0.60, 0.70)	[0.70, 0.80)	[0.80, 0.90)	[0.90, 1.00)
ln(*ANALYSTS*)	1.383	0.877	0.697	0.492	0.387	0.354	0.311	0.266	0.183	0.080
ln(*NREV*)	2.239	1.347	1.060	0.731	0.565	0.519	0.454	0.372	0.239	0.091

Panel B: Pooled sample results after controlling for illiquidity

Sample:	World	World	World	World	World	World
Dependent variable:	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*
ln(*ANALYSTS*)	0.037	0.026	−0.002			
	(2.91)***	(2.22)**	(−0.15)			
ln(*NREV*)				0.039	0.026	0.006
				(5.52)***	(4.04)***	(1.00)
Control for illiquidity						
ZeroReturn [linear]		−1.917			−1.911	
		(−20.94)***			(−20.67)***	
Zero-return bin dummies [nonlinear]			$n = 100$			$n = 100$
Control variables and fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
n	377,598	377,598	377,598	377,598	377,598	377,598
Adjusted R^2	0.446	0.476	0.484	0.446	0.477	0.484

(The table is continued on the next page.)

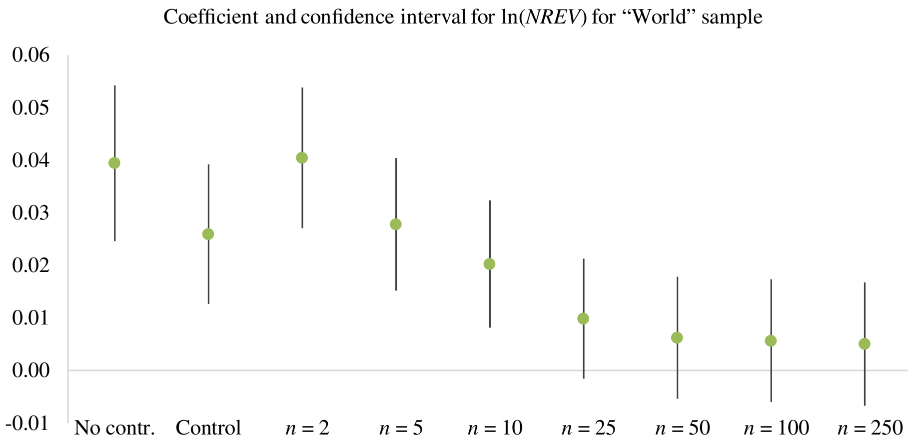
TABLE 5 (continued)

Panel C: Results for country-by-country estimation after controlling for illiquidity

Sample:	World by-country	World by-country	World by-country	World by-country	World by-country	World by-country
Dependent variable:	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*
ln(*ANALYSTS*)	0.135	0.071	0.026			
	(4.44)***	(2.91)***	(0.87)			
ln(*NREV*)				0.088	0.045	0.013
				(4.41)***	(2.89)***	(0.69)
Control for illiquidity						
ZeroReturn [linear]		−2.262			−2.250	
		(−20.78)***			(−20.50)***	
Zero-return bin dummies [nonlinear]			$n = 100$			$n = 100$
Control variables and fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
n	377,598	377,598	377,598	377,598	377,598	377,598
Average adjusted R^2	0.428	0.469	0.483	0.429	0.469	0.483

Notes: NREV equals the number of revisions in analysts' one-year-ahead earnings forecast during the year, as identified in I/B/E/S. *ANALYSTS* captures the average number of one-year-ahead earnings forecasts outstanding during the year, also identified in I/B/E/S. ln(*NREV*) and ln(*ANALYSTS*) are the natural logarithm of one plus *NREV* and one plus *ANALYSTS*, respectively. All other variables are defined as in the notes to Table 3. In panels B and C, we control for illiquidity by including (but not tabulating) 100 dummy variables that capture whether or not an observation falls into one of the 100 bins of zero-return frequency of equal width: [0.00–1.00) with steps of 0.01. Control variables are the control variables included in Table 4. Fixed effects are the country-year (or year) and industry indicator variables included in Table 4. Regression intercepts are included but not tabulated. For the pooled regressions, *t*-statistics (reported in parentheses) are based on robust standard errors adjusted for heteroskedasticity and two-way clustering by firm and year. For the country-level regressions, *t*-statistics are based on the average and SDs of the 50 coefficient estimates. *** and ** indicate statistical significance at the level of 0.01 and 0.05, respectively.

Figure 2 Linear versus nonlinear (dummy variable) controls for illiquidity in the relation between analyst coverage and stock price synchronicity



Notes: This figure presents the coefficient estimates (dots) and associated 95 percent confidence intervals (vertical lines) for the ln(*NREV*) variable for the pooled regressions of the "World" sample in Table 5 using different approaches to controlling for illiquidity. *NREV* is defined as the number of revisions in analysts' one-year-ahead earnings forecast during the year, as identified in I/B/E/S. "Control" refers to the regression including *ZeroReturn* as a control variable. "No contr." indicates the regression is estimated without such control variable. $n = \{2, 5, 10, 25, 50, 100, 250\}$ refers to the number of dummy variables included in the regression to control for illiquidity. For *ZeroReturn*, dummy variables are created based on membership in the n bins in the [0.00, 1.00) interval of zero-return frequency, where each bin has a width of 1.00/*n*. Tests are based on a sample of 377,598 firm-year observations. Confidence intervals are based on robust standard errors adjusted for heteroskedasticity and two-way clustering by firm and year.

potentially eliminate a true relation of interest. In Appendix 2, we use causal diagrams in the form of Directed Acyclic Graphs (DAGs, see, e.g., Morgan and Winship 2015) to illustrate how researchers should carefully consider these possibilities. In this case, we conclude that controlling for illiquidity actually helps, rather than hinders, researchers in identifying the effect of analyst coverage on the relative mix of information reflected in stock prices. The argument that analysts affect the relative amount of market-wide information in prices rests on the proposed mechanism that analysts directly provide this information to market participants. Thus, it does not rely on changes in illiquidity as a channel through which the information is impounded in prices. Controlling for illiquidity helps the researcher to isolate this relation of interest.

Implications for research: Country-level transparency and synchronicity

While the previous tests demonstrate how controlling for measurement bias in R^2 may eliminate the significance of a previously documented relation, measurement bias can also work against a researcher in finding a significant relation where one exists. Therefore, we examine the association of country-level R^2 measures with proxies for corporate transparency in countries before and after controlling for illiquidity. In this setting, prior research unambiguously predicts a negative relation between synchronicity and transparency. However, to the extent that country-level transparency relates negatively to illiquidity, not controlling for illiquidity can cause a positive bias in the coefficient relating country-level transparency to R^2 and hide a true negative effect (see Appendix 2 for a discussion and insights using causal diagrams).

Jin and Myers (2006) examine the relation between synchronicity and a country-level disclosure index. Using a set of 40 countries, they find that opaque countries are associated with more

TABLE 6
Analyst coverage and synchronicity: U.S. results and alternative control for illiquidity

Panel A: U.S. results before and after controlling for illiquidity using zero-return frequency

Sample:	U.S.	U.S.	U.S.	U.S.	U.S.	U.S.
Dependent variable:	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*
ln(*ANALYSTS*)	0.045	0.031	0.017			
	(2.11)**	(1.49)	(0.78)			
ln(*NREV*)				0.073	0.062	0.036
				(7.39)***	(6.40)***	(3.79)***
Control for illiquidity						
ZeroReturn [linear]		−1.477			−1.412	
		(−6.44)***			(−6.31)***	
Zero-return bin dummies [nonlinear]			n = 100			n = 100
Control variables and FEs	Yes	Yes	Yes	Yes	Yes	Yes
n	94,861	94,861	94,861	94,861	94,861	94,861
Adjusted R^2	0.428	0.433	0.451	0.430	0.434	0.451

Panel B: U.S. results before and after controlling for illiquidity using Amihud measure

Sample:	U.S.	U.S.	U.S.	U.S.	U.S.	U.S.
Dependent variable:	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*	*Synch*
ln(*ANALYSTS*)	0.045	−0.002	−0.035			
	(2.11)**	(−0.08)	(−1.56)			
ln(*NREV*)				0.073	0.038	0.009
				(7.39)***	(3.59)***	(0.81)
Control for illiquidity						
ln(*Amihud*) [linear]		−0.195			−0.186	
		(−11.37)***			(−10.55)***	
Amihud portfolio dummies [nonlinear]			n = 100			n = 100
Control variables and FEs	Yes	Yes	Yes	Yes	Yes	Yes
n	94,861	94,861	94,861	94,861	94,861	94,861
Adjusted R^2	0.428	0.439	0.448	0.430	0.439	0.447

Notes: This table presents analyses similar to those in Table 5 for the sample of U.S. firm-year observations. Panel A analyses control for *ZeroReturn*. Panel B analyses replace *ZeroReturn* by the Amihud measure of illiquidity (*Amihud*). *Amihud* is defined as the median of daily estimates of (|*ret*|/prc×vol) within a firm-year based on the CRSP daily stock file, multiplied by 1,000,000. *NREV* equals the number of revisions in analysts' one-year-ahead earnings forecast during the year, as identified in I/B/E/S. *ANALYSTS* captures the average number of one-year-ahead earnings forecasts outstanding during the year, also identified in I/B/E/S. All other variables are defined as in the notes to Table 3. We control for illiquidity either by including *ZeroReturn* (ln(*Amihud*)) as a control variable, or by including, but not tabulating, 100 dummy variables that capture whether or not an observation falls into one of the 100 bins (percentiles) of the *ZeroReturn* (*Amihud*) variable. For the Amihud measure, percentiles are defined based on the pooled sample of 94,861 observations. Control variables are the control variables included in Table 4. Fixed effects are the country-year (or year) and industry indicator variables included in Table 4. Regression intercepts are included but not tabulated. *t*-statistics (reported in parentheses) are based on robust standard errors adjusted for heteroskedasticity and two-way clustering by firm and year. *** and ** indicate statistical significance at the level of 0.01 and 0.05, respectively.

synchronous stock prices. Bartram et al. (2012) expand the list of countries and report the disclosure index for 47 out of our 50 countries. We obtain the disclosure index from their table 1-B. Bushman et al. (2004) investigate a set of characteristics of firms' information environments within different countries and create a financial transparency factor. We obtain this factor score from their appendix B, which reports data for 41 of our countries. The (untabulated) Spearman correlation between the two variables for the 39 countries with overlapping data equals 0.6823.[34]

In Table 7, we examine the relation between country-year average *Synch* and the disclosure index (*Disclosure*) and financial transparency factor (*Factor* 1). Simple regressions suggest that the two variables are not statistically significantly related to country-level synchronicity (columns (1) and (4)), suggesting we cannot conclude that greater transparency is associated with lower synchronicity based on these tests. At the same time, whether we include a control for illiquidity using country-year average *ZeroReturn* or 100 dummy variables based on country-year average *ZeroReturn* bins, the insignificant coefficient becomes more negative and statistically highly significant. Moreover, the statistical significance and explanatory power increase when allowing the relation between illiquidity and synchronicity to be more flexible in the LSDV approach (increases in adjusted R^2 from 0.303 to 0.362 and from 0.223 to 0.303, respectively).[35]

Overall, these tests further illustrate how the illiquidity effect can change inferences drawn about the relation between synchronicity and variables of interest. We find that tests of the relation between a variable of interest and R^2, which can pick up firm-specific information flow (see Table 1), can be biased toward a null result if illiquidity is not controlled for.

6. Summary and conclusions

Prior research uses synchronicity measures (market model R^2s) to capture the extent to which stock prices impound firm-specific information, such as in studies on the capital market consequences of corporate transparency, changes in accounting standards, and transparency regulations, as well as the role of information intermediaries in shaping the mix of information in prices. This paper demonstrates that R^2 measures are predictably biased as a result of variation in stock illiquidity and that traditional remedies in the literature, used to correct market model betas for illiquidity effects, do not effectively fix R^2.

While R^2 can identify firm-specific information flows, we find that measurement bias masks these effects for less liquid stocks. Using a large sample of firm-year observations, we document strong negative relations between illiquidity and R^2 across countries, across firms, and over time. We illustrate the importance of controlling for measures of illiquidity and highlight the consequences of using methods that control for nonlinear confounding effects. We do so by using variables for analyst coverage and country-level corporate transparency for our international sample, and present evidence suggesting that the positive relation between analyst coverage and synchronicity disappears, while the insignificant relation between transparency and synchronicity becomes stronger and significantly negative, after we control for illiquidity.

We conclude that the strong link between illiquidity and R^2 induced by predictable measurement bias provides an explanation for why prior research typically finds low R^2 to be related to relatively weak information environments. Besides elaborating more carefully what it is a

34. The disclosure index (*Disclosure*) from Jin and Myers (2006) is based on data from Global Competitiveness Reports for 1999 and 2000, which contain survey results on the level and effectiveness of financial disclosures in different countries. Bushman et al. (2004) derive their financial transparency factor (*Factor* 1) from a factor analysis of country-level measures of firms' information environments. The factor analysis includes variables that capture the intensity and timeliness of financial disclosures, combined with variables that capture the interpretation and dissemination of these disclosures by analysts and the media.

35. Note that the purpose of our analysis is not to support or reject the hypothesis that better transparency *leads to* lower synchronicity. Our purpose is merely to show that synchronicity in and of itself displays a more consistent relation with country-level transparency after the illiquidity effect is controlled for. We therefore examine the relations in a simple setting without controlling for any other country characteristics.

TABLE 7

Country-level transparency and synchronicity: Controlling for illiquidity in a cross-country setting

	Country-year average *Synch*	Country-year average *Synch*	Country-year average *Synch*	Country-year average *Synch*	Country-year average *Synch*	Country-year average *Synch*
Disclosure	−0.201	−0.280	−0.280			
	(−1.63)	(−3.91)***	(−4.02)***			
Factor1				−0.082	−0.271	−0.287
				(−0.81)	(−3.64)***	(−4.31)***
Control for illiquidity						
ZeroReturn		−2.728			−2.559	
(Country-year average) [linear]		(−8.26)***			(−7.48)***	
Zero-return bin dummies [nonlinear]			$n = 100$			$n = 100$
n	983	983	983	870	870	870
Adjusted R^2	0.038	0.303	0.362	0.007	0.223	0.303

Notes: This table presents analyses of the relation between country-level transparency measures and stock price synchronicity before and after controlling for illiquidity using the frequency of zero returns. All analyses are performed at the country-year level. *Disclosure* is the country-specific disclosure index used in Jin and Myers (2006) and reported in Bartram et al. (2012, Table 1, panel B). This variable is not available for Morocco, Pakistan, and Sri Lanka. *Factor1* is the country-specific financial transparency factor from Bushman et al. (2004, Appendix B). This variable is not available for China, Czech Republic, Egypt, Hungary, Indonesia, Morocco, Portugal, Russian Federation, and Taiwan. Regression *t*-statistics (reported in parentheses) are based on robust standard errors adjusted for heteroskedasticity and clustering by country. Regression intercepts are included but not tabulated. *** indicates statistical significance at the level of 0.01.

synchronicity measure should be capturing conceptually, we advise future researchers to carefully evaluate the sensitivity of their empirical results to (nonlinear) controls for illiquidity.

Appendix 1

Simulations of the effect of non-trading on R^2 measurement

To quantify the effect of non-trading on the measurement of stock price synchronicity, we perform the following simulations. Among the securities in CRSP during 1990–2012 with less than 10 percent of zero-return days and full price data for the calendar year, we randomly draw a security-year.[36] Next, we randomly change a fraction of p nonzero-return trading days into zero-return (non-trading) days. As a result of this induced non-trading, the return for the first subsequent nonzero-return day equals its true return plus the sum of true returns for previous consecutive non-trading days (see section 2).

We repeat the above procedure 99,999 times (with replacement) to obtain a sample of 100,000 simulated observations. We sample the non-trading fraction p based on the frequency distribution in panel B of Table 3 to allow the simulations to be as a realistic as possible. Specifically, based on the observation counts for the nine bins from 0 to 90 percent zero-return frequency, we compute the relative (rounded) frequencies of the nine bins to be 39.2, 22.6, 13.3, 7.7, 5.0, 3.9, 3.2, 2.8, and 2.3 percent, respectively, and we induce non-trading fractions accordingly. For example, for $p = 0.01$, we draw 3,920 observations, and do the same for $p = 0.02$ through $p = 0.10$ such that our first bin contains 39,200 observations (39.2 percent). For $p = 0.11$, we draw 2,260 observations, and so on.

For each security-year drawn we compute $Var[R]$, β, and R^2 based on actual returns, and $Var[R]_{zero}$, β_{zero}, and R^2_{zero} based on the new set of returns with non-trading induced. We compute the ratios $Var[R]_{zero}/Var[R]$, β_{zero}/β, and R^2_{zero}/R^2, respectively. The expected value of β_{zero}/β equals q, while the expected value of R^2_{zero}/R^2 equals $q^2(Var[R]_{zero}/Var[R])$.

Table 8 reveals that non-trading decreases the variance of observed returns in our simulations. This decrease becomes larger with p, but the overall (untabulated) median ratio of return variances is still 0.993.[37] Importantly, for any p, we find that the ratio of return variances is substantially greater than q^2, suggesting the effect of non-trading on R^2 should be negative (see section 2). As expected, the median ratio of betas approximates q. Most importantly, non-trading has a strong negative effect on R^2. For instance, when inducing an average of 25.5 percent of zero returns (a common frequency in many countries), the median ratio of R^2s equals 0.561, suggesting that on average the observed R^2 will be 43.9 percent lower than it would be without the zero returns.

The Spearman correlation between p and R^2_{zero}/R^2 equals −0.6697 for the pooled sample of 100,000 simulated observations. In untabulated tests, we also examine this correlation for subsets of the data with lower non-trading frequencies. Even with relatively low non-trading frequencies, we find that the correlation between p and the bias ratio is substantial. For example, for $p \leq 0.20$ ($p \leq 0.10$), the Spearman correlation still equals −0.468 (−0.348).

We also use simulations to examine the consequences to the market model of common adjustments for infrequent trading. Table 9 presents the effect of non-trading on R^2 measurement using (i) weekly returns, (ii) weekly returns with a lag of market returns in the model, (iii) weekly

36. We use the restriction $p < 0.10$ because the vast majority of securities have at least one zero-return day (especially in earlier years) and it allows us to randomly change up to 90 percent of nonzero-return days for a particular security into zero-return days.

37. This result is different from the predictions in Scholes and Williams (1977) and Lo and MacKinlay (1990) that non-trading increases the variance of observed returns. It is important to note, however, that these predictions are based on the assumption of independent return observations, while the daily return data we use exhibit negative serial correlation (average of −0.032). To assess the influence of such serial correlation, we rerun the simulations by first randomizing the daily returns in the firm-year before inducing the non-trading effects. Doing so, we find the downward bias in return variance largely disappears. We thank an anonymous reviewer for this suggestion.

TABLE 8
Simulation results on biases in synchronicity measurement due to non-trading effects

				$Var[R]_{zero}$ $/Var[R]$	β_{zero}/β		$R^2{}_{zero}/R^2$	
Bin	No. obs.	Avg. p	Avg. q	Median	Expected	Median	Expected	Median
1	38,700	0.055	0.945	0.998	0.945	0.955	0.895	0.916
2	22,700	0.155	0.845	0.991	0.845	0.845	0.721	0.727
3	13,400	0.255	0.745	0.983	0.745	0.737	0.565	0.561
4	7,700	0.355	0.645	0.969	0.645	0.628	0.429	0.418
5	5,000	0.455	0.545	0.961	0.545	0.521	0.309	0.296
6	4,000	0.555	0.445	0.942	0.445	0.410	0.210	0.192
7	3,300	0.655	0.345	0.923	0.345	0.307	0.129	0.117
8	2,800	0.755	0.245	0.880	0.245	0.200	0.068	0.063
9	2,400	0.855	0.145	0.845	0.145	0.099	0.025	0.031
Spearman correlation between p and bias ratio						−0.692		−0.670

Notes: Simulations are based on the following procedure, which we repeat 99,999 times. Among all securities in CRSP during 1990–2012 with less than 10 percent of zero-return days and full price data for the calendar year, we randomly draw (with replacement) a security-year and calculate $Var[R]$, β, R^2, and σ_ε^2 based on the actual daily returns observed in CRSP. Next, for the same security-year, we transform a random fraction of p nonzero-return trading days into non-trading zero-return days. As a result of the non-trading, the return for the first subsequent nonzero-return day equals its true return plus the sum of true returns for previous consecutive non-trading days. We calculate $Var[R]_{zero}$, β_{zero}, R^2_{zero}, and $\sigma^2_{\varepsilon\, zero}$ based on the new set of daily returns with induced non-trading effects. p reflects the proportion of nonzero daily returns that we transform into non-trading days and $q = 1 - p$, where p varies with increments of 0.01 in the interval [0.01, 0.90] and the frequency of each p is based on the relative frequencies of zero returns in the first nine bins presented in panel B of Table 3.

returns with a lead of market returns, (iv) weekly returns with two leads and two lags, (v) and weekly returns with four lags. The fifth adjustment is not common in the R^2 literature, but follows the "price delay" literature, which assumes that most delays in the incorporation of information in prices are resolved within four weeks (Hou and Moskowitz 2005). Importantly, because these estimations differ in their degrees of freedom, we focus on adjusted R^2s.

The results in Table 9 suggest that these approaches partially reduce the absolute magnitude of bias in R^2. At the same time, the negative bias remains nontrivial for larger non-trading frequencies, and the clear negative relation between non-trading frequencies and R^2 bias remains. The Spearman correlations of p with R^2_{zero}/R^2 are −0.399, −0.356, −0.379, −0.321, and −0.321, respectively. These correlations are smaller than those based on daily returns in a simple market model, but their magnitudes suggest that common remedies for infrequent trading do not eliminate the illiquidity effect in R^2 measurement.[38]

Appendix 2

Consequences of empirically controlling for illiquidity

A potential concern with our advice to empirically control for illiquidity is that doing so could eliminate the relation of interest when illiquidity is part of the conceptual mechanism—for example, as

38. Our simulations assume that non-trading occurs at random points in time, while in reality non-trading is more likely clustered in time. When we repeat our simulations and induce a clustering of non-trading over time, we find the inferences of the simulations are unchanged.

TABLE 9
Simulation results based on weekly returns and adjusted market models

		R^2_{zero}/R^2	R^2_{zero}/R^2 (incl. 1 lag)	R^2_{zero}/R^2 (incl. 1 lead)	R^2_{zero}/R^2 (incl. 2 leads/lags)	R^2_{zero}/R^2 (incl. 4 lags)
Bin	Avg. *p*	Median	Median	Median	Median	Median
1	0.055	0.997	0.998	0.997	0.999	0.998
2	0.155	0.939	0.944	0.940	0.946	0.946
3	0.255	0.874	0.886	0.870	0.887	0.891
4	0.355	0.782	0.805	0.783	0.815	0.818
5	0.455	0.671	0.727	0.678	0.736	0.741
6	0.555	0.525	0.620	0.540	0.632	0.650
7	0.655	0.358	0.502	0.374	0.538	0.542
8	0.755	0.172	0.351	0.203	0.418	0.424
9	0.855	0.030	0.176	0.064	0.249	0.263
Spearman correlation		−0.399	−0.356	−0.379	−0.321	−0.321

Notes: See the notes to Table 8 for explanations. Here, R^2 values are obtained from weekly returns and adjusted market models and are based on adjusted R^2 to control for differences in the degrees of freedom across the estimations.

a mediating variable. An example of such a situation occurs with tests of the relation between firm transparency and synchronicity. Below, we illustrate how controlling for illiquidity can provide a lower bound estimate on this relation, rather than eliminate it. While the term "controlling for" typically refers to inclusion of an additional variable in an OLS regression, we define the term more generally to include any parametric or nonparametric approach that makes the statistical relation of interest conditional on a (set of) control variable(s) (Morgan and Winship 2015; Gow et al. 2016), such as matching or LSDV (fixed effects) regressions (Angrist and Pischke 2009).

We present causal diagrams using "DAGs" that capture theoretically predicted causal chains.[39] For ease of exposition and in line with market microstructure theories (e.g., Glosten and Milgrom 1985; Diamond and Verrecchia 1991; Easley and O'Hara 2004), we focus on the following two channels:[40]

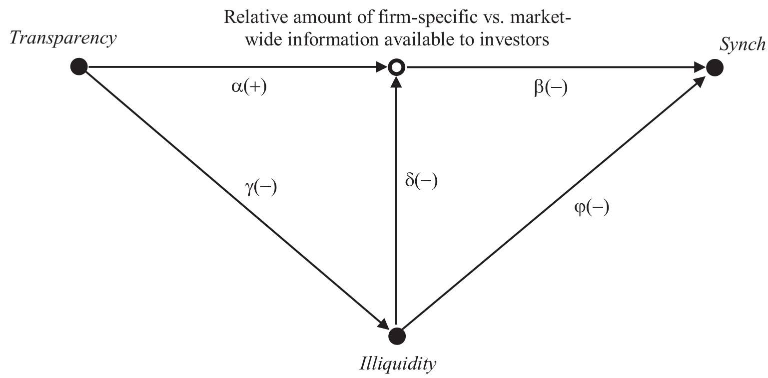
1. *Direct channel*: Greater transparency increases the amount of publicly available firm-specific information; this information is available to all market participants and allows relatively more firm-specific information to be incorporated in prices, thereby increasing stock price informativeness (decreasing synchronicity).
2. *Indirect channel*: Greater transparency reduces information asymmetry between informed investors and uninformed investors as it turns private information into public information; the reduction in adverse selection decreases illiquidity, which reduces the costs of acquiring and trading on firm-specific private information; the increased trading on firm-specific information increases stock price informativeness (decreases synchronicity).

Figure 3 captures these two causal chains. Ignoring for now the additional channel that captures the measurement effect of illiquidity on synchronicity (φ), a regression of the type:[41]

39. Solid dots indicate observable variables and hollow dots indicate unobservable variables. Greek letters and signs indicate theoretically predicted causal effects. See Gow et al. (2016) for more discussion on the use of causal diagrams, and Pearl (1995) or Morgan and Winship (2015) for a formal discussion of using DAGs.

40. By focusing on these two channels, the discussion abstracts away from additional arguments predicting alternative links between transparency and synchronicity. For example, Jin and Myers (2006) link firm-level transparency to managers' rent extraction and the consequences for the level of firm-specific risk reflected in stock prices, while Dasgupta et al. (2010) link transparency to *future* synchronicity.

Figure 3 Causal graph for the effect of public disclosure on synchronicity



$$Synch = a + b\,Transparency + e \tag{9}$$

would yield a coefficient $\hat{b}$ that captures the joint effect of channels #1 and #2, with an expected sign that is negative ($\alpha > 0$, $\gamma\delta > 0$, and $\beta < 0$):

$$E\left[\hat{b}\right] = (\alpha + \gamma\delta)\beta. \tag{10}$$

We next take into account the *measurement* effect of illiquidity, which materializes as a negative relation between illiquidity and synchronicity (φ). With this additional effect, estimation of equation (9) now yields a coefficient $\widehat{b'}$ that captures the following effect:

$$E\left[\widehat{b'}\right] = (\alpha + \gamma\delta)\beta + \gamma\varphi. \tag{11}$$

The expected sign of this coefficient is unclear, because it combines channels #1 and #2 with the measurement effect of illiquidity. While the expected relation between transparency and synchronicity is negative ($(\alpha + \gamma\delta)\beta < 0$), the negative measurement effect of illiquidity (φ) causes the second part of equation (11) to become positive ($\gamma\varphi > 0$). This biases the (predicted negative) coefficient on transparency in equation (9) upward when the measurement effect of illiquidity is not controlled for. This could lead either to a failure to find a true negative relation, or to a conclusion that this relation is positive instead. Now consider the following equation that includes a variable for illiquidity, thereby conditioning the relation between transparency and synchronicity on illiquidity:

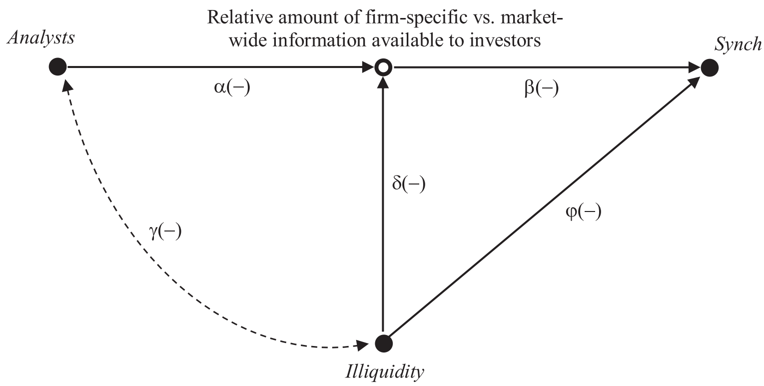
$$Synch = a + c\,Transparency + d\,Illiquidity + e. \tag{12}$$

As stated before, this control can be any nonlinear function of *Illiquidity* (e.g., polynomials, or dummy variables). As Figure 3 illustrates, estimation of equation (12) provides the following coefficients (link γ drops out after including the *Illiquidity* variable):

$$E[\hat{c}] = \alpha\beta, \tag{13a}$$

41. Note that although we model this relation in a linear way, as in a standard OLS regression framework, the model's parameters could also represent more complex and nonlinear relations between the variables.

Figure 4 Causal graph for the effect of analyst coverage on synchronicity



$$E\left[\hat{d}\right] = \delta\beta + \varphi. \tag{13b}$$

Coefficient $\hat{c}$ provides an estimate for the direct information channel (#1). Coefficient $\hat{d}$ combines the measurement effect of illiquidity with the indirect information channel (#2), which means it cannot be interpreted in a meaningful way. Because the expected indirect information effect of illiquidity on synchronicity is positive ($\delta\beta$), while the measurement effect is negative (φ), the expected sign of $\hat{d}$ depends on the strength of the individual effects. The strong negative relation between illiquidity and synchronicity we find in our data suggests the measurement effect substantially outweighs the indirect information effect. Hence, we argue that researchers testing the relation between transparency and synchronicity should base their inferences on tests conditioned on illiquidity. Based on Figure 3, we conclude that the coefficient on transparency measures the direct information effect (channel #1) and can be interpreted as a lower bound estimate of the information effect of transparency on synchronicity. In contrast, the coefficient on a variable of interest cannot be interpreted in a meaningful way *without* controlling for illiquidity. We conclude that unless the expected effect of transparency on synchronicity runs *solely* through indirect channel #2, regressions that control for illiquidity help to identify the effect.

We can also apply this general framework to other settings, such as that of the relation between analyst coverage and synchronicity. If we replace *Transparency* in Figure 3 by analyst coverage in Figure 4 (*Analysts*), link α becomes negative if analysts increase the relative amount of market-wide information, while link γ remains negative but the direction of causality becomes unclear as shown in the curved dashed line (analysts cover firms that are more liquid, while coverage can also lead to greater liquidity). Figure 4 illustrates that in a model that controls for illiquidity, the coefficient on *Analysts* captures the predicted positive relation of interest ($\alpha\beta > 0$). In a model that does not control for illiquidity, the coefficient on *Analysts* is biased with respect to this relation.

While we empirically find that the positive relation between *Analysts* and *Synch* disappears after we control for illiquidity, Figure 4 suggests this null result is unlikely attributable to our over-controlling for illiquidity. That is, controlling for illiquidity in a regression that links *Analysts* and *Synch* leads to a coefficient that captures channel $\alpha\beta$, which is precisely the channel discussed in Piotroski and Roulstone (2004) and Chan and Hameed (2006). For the illiquidity control to incorrectly eliminate a true information-induced positive relation between *Analysts* and *Synch*, this positive relation would have to run solely through the indirect information effect of illiquidity with analyst coverage causing higher liquidity and this leading to *more* public firm-specific information being priced (δ being positive instead of negative).

References

Alford, A. W., and P. G. Berger. 1999. A simultaneous equations analysis of forecast accuracy, analyst following, and trading volume. *Journal of Accounting, Auditing & Finance* 14 (3): 219–40.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5 (1): 31–56.

Amihud, Y., and H. Mendelson. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17 (2): 223–49.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Armstrong, C. S., K. Balakrishnan, and D. Cohen. 2012. Corporate governance and the information environment: Evidence from state antitakeover laws. *Journal of Accounting and Economics* 53 (1–2): 185–204.

Barth, M. E., W. R. Landsman, M. H. Lang, and C. D. Williams. 2018. Effects on comparability and capital market benefits of voluntary IFRS adoption. *Journal of Financial Reporting* 3 (1): 1–22.

Bartram, S. M., G. Brown, and R. M. Stulz. 2012. Why are U.S. stocks more volatile? *Journal of Finance* 67 (4): 1329–70.

Beaver, W. H. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6: 67–92.

Beaver, W. H., M. F. McNichols, and Z. Z. Wang. 2018. The information content of earnings announcements: New insights from intertemporal and cross-sectional behavior. *Review of Accounting Studies* 23 (1): 95–135.

Bekaert, G., C. R. Harvey, and C. Lundblad. 2007. Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies* 20 (6): 1783–831.

Bhattacharya, N., H. Desai, and K. Venkataraman. 2013. Does earnings quality affect information asymmetry? Evidence from trading costs. *Contemporary Accounting Research* 30 (2): 482–516.

Bissessur, S., and A. Hodgson. 2012. Stock market synchronicity—An alternative approach to assessing the information impact of Australian IFRS. *Accounting & Finance* 52 (1): 187–212.

Boudoukh, J., R. Feldman, S. Kogan, and M. Richardson. 2019. Information, trading, and volatility: Evidence from firm-specific news. *Review of Financial Studies* 32 (3): 992–1033.

Bushman, R. M., J. D. Piotroski, and A. J. Smith. 2004. What determines corporate transparency? *Journal of Accounting Research* 42 (2): 207–52.

Callen, J. L., M. Khan, and H. Lu. 2013. Accounting quality, stock price delay, and future stock returns. *Contemporary Accounting Research* 30 (1): 269–95.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu. 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56 (1): 1–43.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay. 1997. *The econometrics of financial markets*. Princeton, NJ: Princeton University Press.

Chan, K., and Y.-C. Chan. 2014. Price informativeness and stock return synchronicity: Evidence from the pricing of seasoned equity offerings. *Journal of Financial Economics* 114 (1): 36–53.

Chan, K., and A. Hameed. 2006. Stock price synchronicity and analyst coverage in emerging markets. *Journal of Financial Economics* 80 (1): 115–47.

Chan, K., A. Hameed, and W. Kang. 2013. Stock price synchronicity and liquidity. *Journal of Financial Markets* 16 (3): 414–38.

Choi, J.-H., S. Choi, L. A. Myers, and D. Ziebart. 2019. Financial statement comparability and the informativeness of stock prices about future earnings. *Contemporary Accounting Research* 36 (1): 389–417.

Christensen, H. B. 2012. Why do firms rarely adopt IFRS voluntarily? Academics find significant benefits and the costs appear to be low. *Review of Accounting Studies* 17 (3): 518–25.

Christensen, H. B., L. Hail, and C. Leuz. 2013. Mandatory IFRS reporting and changes in enforcement. *Journal of Accounting and Economics* 56 (2–3, Supplement 1): 147–77.

Cohen, D. A., A. Dey, and T. Z. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. *The Accounting Review* 83 (3): 757–87.

Cohen, K. J., G. A. Hawawini, S. F. Maier, R. A. Schwartz, and D. K. Whitcomb. 1983. Friction in the trading process and the estimation of systematic risk. *Journal of Financial Economics* 12 (2): 263–78.

Coller, M., and T. L. Yohn. 1997. Management forecasts and information asymmetry: An examination of bid-ask spreads. *Journal of Accounting Research* 35 (2): 181–91.

Crawford, S. S., D. T. Roulstone, and E. C. So. 2012. Analyst initiations of coverage and stock return synchronicity. *The Accounting Review* 87 (5): 1527–53.

Dasgupta, S., J. Gan, and N. Gao. 2010. Transparency, price informativeness, and stock return synchronicity: Theory and evidence. *Journal of Financial and Quantitative Analysis* 45 (5): 1189–220.

Daske, H., L. Hail, C. Leuz, and R. Verdi. 2008. Mandatory IFRS reporting around the world: Early evidence on the economic consequences. *Journal of Accounting Research* 46 (5): 1085–142.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance* 46 (4): 1325–59.

Dimson, E. 1979. Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics* 7 (2): 197–226.

Dimson, E., and P. R. Marsh. 1983. The stability of UK risk measures and the problem of thin trading. *Journal of Finance* 38 (3): 753–83.

Dong, Y., O. Z. Li, Y. Lin, and C. Ni. 2016. Does information-processing cost affect firm-specific information acquisition? Evidence from XBRL adoption. *Journal of Financial and Quantitative Analysis* 51 (2): 435–62.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41 (5): 797–836.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59 (4): 1553–83.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–36.

Fernandes, N., and M. A. Ferreira. 2008. Does international cross-listing improve the information environment? *Journal of Financial Economics* 88 (2): 216–44.

Ferreira, M. A., and P. A. Laux. 2007. Corporate governance, idiosyncratic risk, and information flow. *Journal of Finance* 62 (2): 951–89.

Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14 (1): 71–100.

Gordon, E. A., A. Greiner, M. J. Kohlbeck, S. Lin, and H. Skaife. 2013. Challenges and opportunities in cross-country accounting research. *Accounting Horizons* 27 (1): 141–54.

Gormley, T. A., and D. A. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27 (2): 617–61.

Gow, I. D., D. F. Larcker, and P. C. Reiss. 2016. Causal inference in accounting research. *Journal of Accounting Research* 54 (2): 477–523.

Grewal, J., C. Hauptmann, and G. Serafeim. 2017. Material sustainability information and stock price informativeness. Working paper, Harvard Business School and Oxford University, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2966144.

Griffin, J. M., P. J. Kelly, and F. Nardari. 2010. Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. *Review of Financial Studies* 23 (8): 3225–77.

Grossman, S. J., and J. E. Stiglitz. 1976. Information and competitive price systems. *American Economic Review* 66 (2): 246–53.

Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70 (3): 393–408.

Gul, F. A., J.-B. Kim, and A. A. Qiu. 2010. Ownership concentration, foreign shareholding, audit quality, and stock price synchronicity: Evidence from China. *Journal of Financial Economics* 95 (3): 425–42.

Gul, F. A., B. Srinidhi, and A. C. Ng. 2011. Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics* 51 (3): 314–38.

Haggard, K. S., X. Martin, and R. Pereira. 2008. Does voluntary disclosure improve stock price informativeness? *Financial Management* 37 (4): 747–68.

Han, Y., and D. Lesmond. 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. *Review of Financial Studies* 24 (5): 1590–629.

Hou, K., and T. J. Moskowitz. 2005. Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies* 18 (3): 981–1020.

Hutton, A. P., A. J. Marcus, and H. Tehranian. 2009. Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics* 94 (1): 67–86.

Irvine, P. J., and J. Pontiff. 2009. Idiosyncratic return volatility, cash flows, and product market competition. *Review of Financial Studies* 22 (3): 1149–77.

Israeli, D., C. M. C. Lee, and S. A. Sridharan. 2017. Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies* 22 (3): 1048–83.

Jin, L., and S. C. Myers. 2006. R^2 around the world: New theory and new tests. *Journal of Financial Economics* 79 (2): 257–92.

Kelly, B., and A. Ljungqvist. 2012. Testing asymmetric-information asset pricing models. *Review of Financial Studies* 25 (5): 1366–413.

Kelly, P. J. 2014. Information efficiency and firm-specific return variation. *Quarterly Journal of Finance* 4 (4): 1450018.

Kim, J.-B., and H. Shi. 2012. IFRS reporting, firm-specific information flows, and institutional environments: International evidence. *Review of Accounting Studies* 17 (3): 474–517.

Kim, Y., H. Li, and S. Li. 2012. Does eliminating the form 20-F reconciliation from IFRS to U.S. GAAP have capital market consequences? *Journal of Accounting and Economics* 53 (1–2): 249–70.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53 (6): 1315–35.

Lang, M., K. V. Lins, and M. Maffett. 2012. Transparency, liquidity, and valuation: International evidence on when transparency matters most. *Journal of Accounting Research* 50 (3): 729–74.

Lesmond, D. A. 2005. Liquidity of emerging markets. *Journal of Financial Economics* 77 (2): 411–52.

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka. 1999. A new estimate of transaction costs. *Review of Financial Studies* 12 (5): 1113–41.

Leuz, C., and R. E. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91–124.

Li, B., S. Rajgopal, and M. Venkatachalam. 2014. R^2 and idiosyncratic risk are not interchangeable. *The Accounting Review* 89 (6): 2261–95.

Lo, A. W., and C. A. MacKinlay. 1990. An econometric analysis of nonsynchronous trading. *Journal of Econometrics* 45 (1–2): 181–211.

Morck, R., B. Yeung, and W. Yu. 2000. The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58 (1–2): 215–60.

Morck, R., B. Yeung, and W. Yu. 2013. R^2 and the economy. *Annual Review of Financial Economics* 5 (1): 143–66.

Morgan, S. L., and C. Winship. 2015. *Counterfactuals and causal inference*, 2nd ed. New York: Cambridge University Press.

Muslu, V., M. Rebello, and Y. Xu. 2014. Sell-side analyst research and stock comovement. *Journal of Accounting Research* 52 (4): 911–54.

Pearl, J. 1995. Causal diagrams for empirical research. *Biometrika* 82 (4): 669–88.

Peterson, K., R. Schmardebeck, and T. J. Wilks. 2015. The earnings quality and information processing effects of accounting consistency. *The Accounting Review* 90 (6): 2483–514.

Piotroski, J. D., and D. T. Roulstone. 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The Accounting Review* 79 (4): 1119–51.

Piotroski, J. D., T. J. Wong, and T. Zhang. 2015. Political incentives to suppress negative information: Evidence from Chinese listed firms. *Journal of Accounting Research* 53 (2): 405–59.

Roll, R. 1988. R^2. *Journal of Finance* 43 (3): 541–66.

Roulstone, D. T. 2003. Analyst following and market liquidity. *Contemporary Accounting Research* 20 (3): 552–78.

Scholes, M., and J. Williams. 1977. Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5 (3): 309–27.

Skaife, H. A., D. Veenman, and D. Wangerin. 2013. Internal control over financial reporting and managerial rent extraction: Evidence from the profitability of insider trading. *Journal of Accounting and Economics* 55 (1): 91–110.

Teoh, S. H., Y. G. Yang, and Y. Zhang. 2009. *R*-square and market efficiency. Working paper.

Veldkamp, L. L. 2006. Information markets and the comovement of asset prices. *Review of Economic Studies* 73 (3): 823–45.

Verrecchia, R. E. 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50 (6): 1415–30.

Wang, J. W., and W. W. Yu. 2015. The information content of stock prices, legal environments, and accounting standards: International evidence. *European Accounting Review* 24 (3): 471–93.

Wang, S. 2019. Informational environments and the relative information content of analyst recommendations and insider trades. *Accounting, Organizations and Society* 72: 61–73.

Watanabe, A., Y. Xu, T. Yao, and T. Yu. 2013. The asset growth effect: Insights from international equity markets. *Journal of Financial Economics* 108 (2): 529–63.

Xing, X., and R. Anderson. 2011. Stock price synchronicity and public firm-specific information. *Journal of Financial Markets* 14 (2): 259–76.