International Review of Economics and Finance 39 (2015) 149–161

Contents lists available at ScienceDirect

International Review of Economics and Finance

journal homepage: www.elsevier.com/locate/iref

Does PIN measure information? Informed trading effects on returns and liquidity in six emerging markets

Diego A. Agudelo *, Santiago Giraldo [1], Edwin Villarraga [1]

Department of Finance, Universidad EAFIT, Carrera 49 No. 7 Sur 50, Medellín, Colombia

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form 4 April 2015
Accepted 6 April 2015
Available online 16 April 2015

JEL classification:
G10
G15
G19

Keywords:
Informed trading
Liquidity
PIN
Emerging markets
Market microstructure

ABSTRACT

Market microstructure models imply that informed trading reduces liquidity and moves prices in the direction of the information. We test this implication using the dynamic PIN model (Easley, Engle, O'Hara and Wu 2008) as a time-varying measure of informed trading in the six largest Latin America stock markets. Under alternative specifications and robustness tests, the results suggest that signed dynamic PIN is related to returns, as a proxy for information asymmetry rather than just liquidity effects. These results contribute to the ongoing discussion on whether PIN is a valid informed trading measure, and to a better understanding of price formation in emerging markets.

1. Introduction

Classical microstructure models imply that information asymmetry affects prices and liquidity on financial markets (Kyle, 1985; Glosten & Milgrom, 1985; Easley and O'Hara, 1992). These models argue that informed traders improve market efficiency by exploiting their informational advantage and thus contribute to a more rapid adjustment of prices towards fundamental values. In turn, the liquidity provider faces adverse selection by having to trade with unidentified informed traders hidden among many uninformed traders. The higher the probability of informed trading, the larger the transaction costs and the lower the liquidity. All in all, information asymmetry allows informed traders to earn extra returns at the expense of uninformed traders. Informed trading, in turn, should cause prices to better reflect fundamentals and force liquidity providers to increase trading costs.

Empirical studies of information asymmetry in financial markets hinge critically on a valid measure of informed trading. Easley, Kiefer, O'Hara, and Paperman (1996), and Easley, Kiefer, and O'Hara (1997) present the probability of informed trading (PIN) as a reliable proxy of information asymmetry, based on the assumption that informed traders cause an important part of the observed order imbalance. Using the data on directional individual trades, the PIN model estimates the probabilities of informed and uninformed trading using as inputs the total number of trades and the order imbalance.[2] Those early PIN models yield what we

* Corresponding author. Tel.: +57 4 2619500x9719.
E-mail addresses: dagudelo@eafit.edu.co (D.A. Agudelo), agogiraldo@hotmail.com (S. Giraldo), edwin.villarraga@gmail.com (E. Villarraga).
[1] Tel.: +57 4 2619500x9719.
[2] Theoretical support for the PIN model can be found in Copeland and Galai (1983), Kyle (1985) and Glosten and Milgrom (1985).

http://dx.doi.org/10.1016/j.iref.2015.04.002

call a "static PIN", since they assume constant arrival rates of informed and uninformed trades and typically are estimated in a stock-quarter basis. A numerous literature has used alternative varieties of the static PIN as a measure of information asymmetry, for example Easley, Hvidkjaer, and O'Hara (2002), Chung, Li, and McInish (2005), Vega (2006), and more recently, Chung, Elder, and Kim (2010), Chen and Zhao (2012), Lin, Lee, and Wang (2013), Sankaraguruswamy, Shen, and Yamada (2013) and Chang and Lin (2015). By construction the "static" PIN is limited to measure cross-sectional variation of informed trading, rather than time-series effects. As an alternative, Easley, Engle, O'Hara, and Wu (2008) (henceforth EEOW) propose a dynamic model allowing for time-varying arrival of informed and uninformed trades. The authors present evidence of a direct relationship between dynamic PIN and liquidity for a sample of 16 US stocks.

This paper presents evidence that informed trading, estimated by the dynamic PIN, causes prices to move on the direction of the information and simultaneously reduces liquidity, as predicted by the market microstructure literature. To estimate informed trading we use the dynamic PIN model of EEOW (2008) which, to the extent of our knowledge, has only been run on the options markets (Engle & Neri, 2010). Specifically, we test whether the dynamic PIN is related to liquidity and returns on the six largest Latin American stock markets as predicted by the theory. We know of no previous test on the ability of the dynamic PIN measure to predict liquidity and returns in a wide sample of stocks.

The contribution of this paper to the literature is twofold. First, by focusing on six emerging markets instead of the US, this study provides an out-of-sample test of the theoretical relation between informed trading and liquidity and returns. Whereas abundant evidence have been provided in US markets on the relation between information and liquidity (e.g. Chung et al., 2005; Lei & Wu, 2005), not much has been provided on the times series effect of informed trading on realized returns. Second, this paper contributes to the ongoing debate on whether the family of PIN models renders a valid measure of information asymmetry. Some evidence has cast doubt on the validity of PIN. Using the Static PIN model on T-Bills, Akay, Cyree, Griffiths, and Winters (2012) argue that PIN measures trading clusters rather than information. Aktas, de Bodt, Declerck, and Van Oppens (2007) report that static PIN is unable to detect information leaking around M&A announcements. However, those results could be explained by the inability of Static PIN to detect short-term variations in informed trading. In turn, Duarte and Young (2009) and Lai, Ng, and Zhang (2014), studying samples on US and on 47 international markets respectively, fail to find a relation between the static PIN and the cross-section distribution of returns. However, the absence of a cross sectional relationship between PIN and returns doesn't invalid PIN as an information measure. The effect of idiosyncratic information on prices is expected to be diversified away and thus should not be a priced risk factor.

Our results are supportive of the PIN as a valid informed trading measure, based on two critical differences with previous research. First, we use the dynamic PIN model, which, unlike the Static PIN, is able to detect changes on information asymmetry over time. Second, we provide evidence that dynamic PIN has a permanent effect on prices at daily frequency, only slightly reversed at the next day, which cannot be explained if PIN is just detecting liquidity effects not related to information. This permanent effect of the dynamic PIN on prices is robust under several alternative specifications that take care of three confounding effects: differential effects of PIN on returns on individual stocks, the endogeneity between daily returns and informed trading, and the bid-ask bounce effects on daily returns.

The group of six Latin American emerging markets is an interesting and barely explored object for market microstructure, for their wide variety of size, liquidity and stages of development. The liberalization of Latin American emerging markets in the late 80's and early 90's, as well as their impressive performance in the 2000's, has heralded their increasing role in the world financial system. However, concerns remain about their liquidity, institutional design, governance and efficiency (Kearney, 2012).

Market microstructure studies have been mostly conducted in individual exchanges of US and other G7 countries, without much comparison between international markets. A direct precedent of the current study is Cruces and Kawamura (2005) who estimate the static PIN for seven Latin-American stock markets, finding a cross-sectional relationship between the quality of corporate governance and the average PIN across countries. Moreover, two recent studies have used PIN as a proxy of informed trading in Brazil (Barbedo, Camilo, Pereira, & Leal, 2010; Martins, Paulo, & Albuquerque, 2013). Additionally, Villarraga, Giraldo, and Agudelo (2012) study the distribution of dynamic PIN in the same sample of six emerging markets, focusing on the relation with trading activity, size and day-of-the-week. Two other precedents are Lesmond (2005) who conducts a comprehensive study of liquidity in 31 emerging markets, in quarterly frequency and Bekaert, Harvey, and Lundblad (2007), who test whether liquidity is a priced factor in a set of 19 emerging markets, both studies using liquidity proxies.

The rest of this paper is organized as follows: Section 2 provides the background for the dynamic PIN model and the theoretical relationship between information asymmetry and asset liquidity and returns. Section 3 describes the methodology, providing details on the estimation of the dynamic PIN model and the econometric approach. Section 4 presents and discusses the results found for the six Latin-American stock markets. Finally, Section 5 concludes.

2. Background

2.1. Estimating the dynamic probability of informed trading

The static PIN model describes the arrival of informed and uninformed traders to a market, where a designated market maker provides liquidity (Easley et al., 2002; Easley, Kiefer, O'Hara and Paperman, 1996; Easley et al., 1997; Easley & O'Hara, 1992). Several studies have extended this framework by allowing the rates of arrival of both types of traders to vary over time. Lei and Wu (2005) propose a Markov Switching model of informed and uninformed arrival, resulting in a time-varying PIN model.

Tay, Ting, Tse, and Warachka (2009) present a dynamic PIN model based on asymmetric autoregressive conditional duration that allows for a joint modeling of the duration and direction of trades, enabling an intraday PIN estimation. Easley, Lopez de Prado, and

O'Hara (2012) develop a volume-synchronized probability of informed trading (VPIN) and use it to detect "toxic" order flow in high frequency data, i.e. the trading volume that adversely selects market markers.

EEOW (2008) develops a PIN model with dynamic rates of arrival of informed and uninformed traders conditional on their past values. This way, their time-varying PIN can be estimated from the previous history of trades, as described in Section 3.1. The time-varying arrival of informed and uninformed trades is represented in a bivariate time-series model with auto- and cross-correlations. Autocorrelation of uninformed trades can be understood as herding effects (Lee, Liu, Roll, & Subrahmanyan, 2004; Scharfstein & Stein, 1990) or behavioral biases (Barber, Odean, & Zhu, 2009), autocorrelation of informed trades, as the splitting of large orders (Harris, 2003; Kyle, 1985) Using this dynamic PIN, they found an increase (reduction) of informed trading before (after) earnings announcements.

Duarte and Young (2009) present an extension of the static PIN model that allows for days of high balanced buy and sell activity (due to symmetric order-flow shocks, not to information), as well as different rates of arrival of buys and sells by both informed and uninformed traders. As a result, they derive two static measures: Adj_PIN that better measures informed trading, and PSOS, probability of symmetric order-flow shocks, which measures non-informational liquidity shocks. The authors provide evidence of a significant relation of the cross-sectional variation of returns with PIN, but not with Adj_PIN. Instead they show that the relation between PIN and returns is driven by the PSOS, not by informed trading.

2.2. Informed trading effects on liquidity and returns

Market microstructure theory poses that informed trading should reduce liquidity. The market maker widens the bid-ask spread and/or increases the cost of large trades, anticipating the adverse selection problem she faces for trading with informed traders (Hasbrouck, 1991a; Kyle, 1985). In this regard, Easley, Hvidkjaer, and O'Hara (2010) derive a theoretical relationship between PIN and bid-ask spreads, and Easley et al. (1997, 1996) provide empirical evidence of this relation in US stocks. Chung et al. (2005), using the PIN measure, provide evidence that larger information asymmetry increases the price impact of trades. Finally, both Lei and Wu (2005) and EEOW (2008) go further, by using the time varying PIN estimations to predict bid-ask spreads.

Market microstructure theory also implies that informed trading should move prices in the direction of the information (Easley & O'Hara, 1992; Glosten & Milgrom, 1985; Kyle, 1985). However, to the extent of our knowledge, few evidence has been provided on the short term relation between informed trading and prices. Two streams of earlier research have come close to this point. First, Hasbrouck (1991b) provides empirical evidence on the effect of trading on prices ("trade informativeness"). Second, some studies show that informed trading, measured by PIN, is a priced risk factor in the cross-sectional distribution of returns (e.g. Easley et al., 2010), but there is still dispute whether information or liquidity effects are driving those results (Akay et al., 2012; Duarte & Young, 2009).

The two most closely related studies to the present work are Duarte and Young (2009) and Lai et al. (2014). Both papers provide evidence against the static PIN being a priced factor in a cross-section of expected returns. Lai et al. (2014) estimate a static PIN (Easley et al., 2002) in a sample of 30,095 firms from 47 countries failing to find PIN as a priced factor in a cross-section of average returns. In turn, as described above, Duarte and Young (2009) propose an extension of the static PIN model that provides two static measures: Adj_PIN, a refined measure of informed trading, and PSOS, a measure of liquidity shocks. They show that the relation between PIN and returns is driven by PSOS, which measures non-informational liquidity shocks, and not by Adj_PIN, which captures informed trading.

This paper is different in two ways. First, we are using the Dynamic PIN model that not only is time-varying but also overcomes several of the limitations Duarte and Young (2009) find in the static PIN model. Specifically, the dynamic modeling of the arrival of informed and uniformed traders allows for the symmetric order-flows accounted for in the model of Duarte and Young (2009). Second, we test the time-series effect of the signed Dynamic PIN on returns in the short run, as implied by asymmetric information models of liquidity. We argue that this is a more direct test of the informational component of PIN, than a cross-sectional regression of PIN against average returns. Moreover, our finding of a short-term relation between informed trading and returns, controlling for market return, doesn't contradict the null cross-sectional relation reported by Duarte and Young (2009) and Lai et al. (2014). Once controlling for market returns, idiosyncratic information should move stock prices but not be related to the cross sectional distribution of returns, since it can be either diversified away or subsumed by a systemic risk factor.

3. Methodology and data

3.1. Dynamic PIN

Easley and O'Hara (1992), Easley, Kiefer, O'Hara and Paperman (1996) and Easley et al. (1997) static PIN models start with a market that includes a competitive market maker who trades a risky asset with both informed and uninformed traders. Information events happen in trading days with a probability α. The event conveys negative information with a probability δ, and a positive with probability $(1 - \delta)$. Only in days with information events, informed traders' orders enter the market following a Poisson process with arrival rate μ, only in days with information, always trading on the "right" side of the market: any informed trade is a sell in a negative information day, and a buy in a positive one. In turn, uninformed traders' orders enter the market following a Poisson process with an arrival rate ε, independent on the day information, and are equally likely to be a buy or a sale. Fig. 1 summarizes the described information and trade arrival.

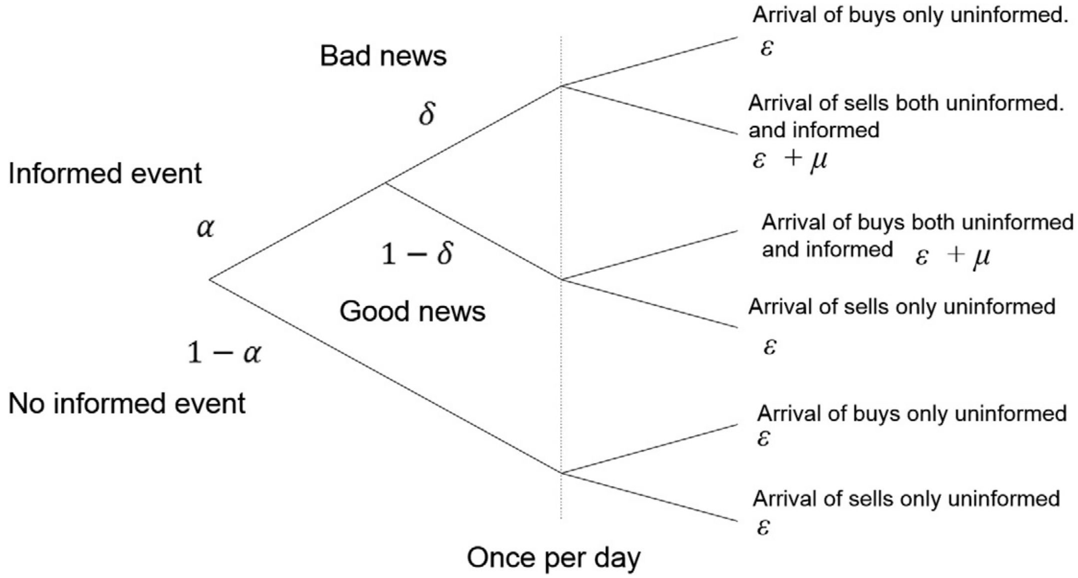

Fig. 1. PIN arrival process.
Source: Easley et al. (1996).

In such a model, the probability of observing a number of B buys and S sells in a given day is as follows:

$$\begin{aligned} Pr[y_t=(B,S)] &= \alpha(1-\delta)e^{-(\mu+2\varepsilon)}\frac{(\mu+\varepsilon)^B(\varepsilon)^S}{B!S!}+\alpha\delta e^{-(\mu+2\varepsilon)}\frac{(\mu+\varepsilon)^S(\varepsilon)^B}{B!S!} \\ &+ \quad (1-\alpha)e^{-2\varepsilon}\frac{(\varepsilon)^{B+S}}{B!S!} \end{aligned} \tag{1}$$

where y_t is the vector of observations (number of buys and sales) in day t. The probability in Eq. (1) is a function of three Poisson process probabilities, weighted by the probability of being in a positive information day $\alpha(1-\delta)$, a negative one $(\alpha\delta)$ and a no information day $(1-\alpha)$. This is a static model in the sense that the arrival of informed and uninformed traders, conditional to the information events, is constant from day to day.

Whereas the static PIN model assumes constant arrival rates for informed and uninformed traders (μ,ε) a more realistic approach provides for agents to continuously update their estimations on the arrival rates, based on the observed trading activity. In that regard, EEWO (2008) offers a methodology to estimate those dynamic rates, allowing the probabilities of buys and sells to vary over time conditional to the predictions in the arrival rates, yielding as a result a dynamic measure of PIN.

According to EEWO (2008), defining the total number of trades as the sum of the number of buys and sells ($TT=S+B$), and the daily order imbalance, as the difference ($K=S-B$), the expected number of trades in a day and the expected value of the order imbalance are as follows:

$$\begin{aligned} &E[TT]=\alpha(1-\delta)(2\varepsilon+\mu)+\alpha\delta(2\varepsilon+\mu)+(1-\alpha)(2\varepsilon)=\alpha\mu+2\varepsilon \\ &E[K]=\alpha\mu(2\delta-1). \end{aligned} \tag{2}$$

Thus, the average of the order imbalance is related to the informed trading arrival. However, a more informative signal is the absolute value of the order imbalance, which is approximated as $E[|K|]\doteq\alpha\mu$. Hence, agents can use those relations to estimate the arrival rates of the two types of traders.

Using the vector $\psi_t=[\alpha\mu_t\,,2\varepsilon_t\,]^T$ to symbolize the two dynamic arrival rates, the deterministic trend is eliminated as $\tilde{\psi}_{it}=\psi_{it}*e^{-g_it}, i=1,2$, where the vector $g=[g_1\,,g_2]^T$ is the rate of growth of the two components of ψ. Then, the untrended arrival rate vector is specified to follow a bivariate autoregressive process, similar to a GARCH model:

$$\tilde{\psi}_t=\omega+\sum_{k=1}^{p}\Phi_k\tilde{\psi}_{t-k}+\sum_{j=0}^{q-1}\Gamma_j\tilde{Z}_{t-j} \tag{3}$$

where: $\tilde{\psi}_t$ is the predicted arrival vector in $t+1$ (detrended) as predicted in t. $\tilde{Z}_{it}=Z_{it}*e^{-g_it}$, i = 1, 2, similarly detrended as above from $Z_t=[|K|, TT_t-|K|]^T$.

To estimate the model EEWO set $p=q=1$, in the original trended variables.

$$\psi_t=\omega\odot e^{gt}+\Phi\left[\psi_{t-1}\odot e^{g}\right]+\Gamma Z_t$$

where $\odot$ denotes the Adamant product. α is assumed to be constant over time, to be able to extract μ from the estimated $\alpha\mu$. Φ and Γ are the 2 × 2 matrices that determine the auto- and cross-correlations of the uninformed and informed rates of arrival. The estimated variables allow the estimation of the probability of observing a given number of buys and sells in a given day, (B_t, S_t) as follows:

$$\begin{aligned} Pr[y_t = (B_t, S_t)|\mathcal{F}_{t-1}] \\ &= \alpha(1-\delta)e^{-(\mu_{t-1}+2\varepsilon_{t-1})}\frac{(\mu_{t-1}+\varepsilon_{t-1})^{B_t}(\varepsilon_{t-1})^{S_t}}{B_t!S_t!} \\ &\quad + \alpha\delta e^{-(\mu_{t-1}+2\varepsilon_{t-1})}\frac{(\mu_{t-1}+\varepsilon_{t-1})^{S_t}(\varepsilon_{t-1})^{B_t}}{B_t!S_t!} + (1-\alpha)e^{-2\varepsilon_{t-1}}\frac{(\varepsilon_{t-1})^{B_t+S_t}}{B_t!S_t!}. \end{aligned}$$

Taking together several days, we have the following aggregate log likelihood function.

$$\sum_{t=1}^{T} \ln Pr[y_t = (B_t, S_t)|\mathcal{F}_{t-1}]$$

Finding the Maximum Likelihood an estimation of the parameters of the model: δ, α, ω, g, Φ, and Γ are obtained. Finally, once the time-varying arrival rates are estimated, the dynamic probability of informed trading, PIN_t, is calculated as follows:

$$PIN_t = \frac{\alpha\mu_t}{\alpha\mu_t + 2\varepsilon_t}.$$

3.2. Estimating the dynamic PIN and liquidity measures

We estimate the dynamic PIN for the six largest Latin American stock markets: Argentina, Brazil, Chile, Colombia, Mexico and Peru. This sample provides an out-of-sample test of short term effects of PIN on returns, as most of the literature has focused on US markets. Besides, it provides variety in terms of number of size and development. Specifically, among the universe of emerging markets reported to the World Federation of Exchanges (2010), Brazil holds the fourth largest emerging stock market, Mexico's is moderately large, Chile's is around the median, both Colombia's and Peru's are small and Argentina has the second smallest. Finally, running independent regressions for each of six emerging markets, as shown below, makes sure our results are general enough.

From Bloomberg we retrieved the tick-by-tick data on quotes (bid and ask prices), transaction prices and traded volume, for the period August 2, 2010 to March 4, 2011.[3] We used several filters to select the stocks. Starting from a total of 1,073 stocks traded in the period, we chose 582 with an average daily trading value over US$ 10,000. We further restrict the sample to the 343 stocks that traded in more than 90% of the trading days to be able to estimate the dynamic PIN model. This final set includes about 88% of the total trading value of the sample, and covering at least 70% in each country.

From Bloomberg tick-by-tick data we find the outstanding bid and ask quotes for each trade in the continuous electronic market, dropping trades and quotes from open or closing call markets. Using the Lee-Ready algorithm (Lee & Ready, 1991), each trade is classified as a buy or a sell.[4] In a stock-day basis, we calculated the number of balanced trades ($TT = S + B$) as well as the trade imbalance ($K = S - B$). For each stock in the sample we estimate the parameters of the dynamic PIN by Maximum Likelihood, as described above.[5] The final result is the estimation of the dynamic PIN measure in a stock-day basis.

From the intraday data we estimate two liquidity measures: the quoted bid-ask spread and the effective bid-ask spread on a stock-day basis (Chordia, Roll, & Subrahmanyam, 2002; Goyenko, Holden, & Trzcinka, 2009). The quoted spread is estimated as the average of the logarithmic difference between the bid and ask prices at the end of each 5-min interval t:

$$quoted\_spread_t = \ln(Bid_t) - \ln(Ask_t). \quad (4)$$

The effective spread is calculated using the transaction price P_k for the k-th trade and the midpoint price M_k, namely the average between the prevailing bid and ask prices, as follows:

$$effective\_spread_k = 2|\ln(P_k) - \ln(M_k)|. \quad (5)$$

[3] For each country in the sample, the period straddles on the last part of the 2010 bull market, and the first part of the 2011 bear market. Besides, it is a period of moderate and decreasing worldwide volatility as given by the evolution of the VIX index.

[4] The six Latin American stock markets are order-driven markets, for the most part lacking a designated market maker. However, Lee and Ready (1991) classification procedure has been formerly used in order-driven markets as Euronext (Aktas et al., 2007).

[5] We use MATLAB, starting each optimization with 120.000 different seeds to better estimate the global maximum inside the feasible optimization region (Yan & Zhang, 2012).

Table 1
Summary statistics of the dynamic PIN model and liquidity measures.

Country	No. of stocks	No. of stock-days	PIN				Quoted bid-ask spread			Effective bid-ask spread			Value of trading (US$ Million) − 2011	Domestic market capitalization (US$ Million) − 2011
			Average	10% and 90% percentile		Median interquartile range per stock	Average	10% and 90% percentile		Average	10% and 90% percentile			
Argentina	12	1,607	0.325	0.089	0.642	0.263	2.72%	0.79%	5.32%	1.48%	0.45%	2.82%	3,815	63,910
Brazil	173	22,666	0.254	0.028	0.566	0.173	0.85%	0.07%	2.00%	0.98%	0.07%	2.41%	868,094	1,171,625
Chile	59	7,887	0.281	0.054	0.592	0.215	1.30%	0.32%	2.77%	0.98%	0.25%	1.94%	53,308	341,799
Colombia	24	3,209	0.254	0.047	0.544	0.205	1.31%	0.39%	2.48%	0.66%	0.19%	1.30%	28,127	208,502
Mexico	53	6,859	0.235	0.030	0.505	0.155	0.76%	0.13%	1.73%	0.60%	0.08%	1.11%	120,064	454,345
Peru	22	2,854	0.339	0.076	0.746	0.243	2.08%	0.69%	3.93%	1.55%	0.47%	3.10%	5,010	103,347

This table reports the summary statistics of the dynamic Probability of Informed Trading (PIN) estimated as described on EEOW (2008), and two intraday liquidity measures, the quoted bid-ask spread and the effective bid-ask spread. The sample comes from six Latin American stock markets, from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. For each variable the average and 10% and 90% percentiles of the stock-day distribution per country are presented. For the PIN, the median of the interquartile range per stock is reported. Value of trading and Domestic market capitalization are taken at the end of 2011 as reported by the Iberoamerican Federation of Stock Exchanges.

3.3. Summary statistics

Table 1 summarizes the estimation of the dynamic PIN and the liquidity measures. Average PIN and liquidity measures differ among the studied stock markets, according to their trading activity and market capitalization. A *t*-test confirms that the largest and more active markets, Brazil and Mexico, exhibit a lower degree of asymmetric information and smaller bid-ask spreads than the others. In contrast, Argentina and Peru, the markets with the lower market cap, trading value and number of firms in the sample, exhibit the highest degree of asymmetric information and the largest bid-ask spreads. Lower degree of asymmetric information as measured by PIN has been related to larger, more transparent and actively traded markets both in studies for worldwide markets (Lai et al., 2014); and for Latin American stock markets (Cruces & Kawamura, 2005).[6] Besides, we calculate the dynamic PIN interquartile range per stock and report its median per country in Table 1. Median ranges of around 0.2 in each country indicate a substantial variation over time of dynamic PINs on most stocks, a clear gain over the static PIN measure.

We also check for a weekly pattern of information asymmetry (Table 2). We find evidence of the U-pattern on the quoted bid-ask reported by Chordia et al. (2002) in the US. Table 2 also shows a U-pattern on the information asymmetry, statistically significant. A more in-depth study of the distribution of the dynamic PIN an its parameters can be found in Villarraga et al. (2012).

To better describe the dynamics of PIN over time we run a panel Tobit model of the daily change of PIN on a set of stock-specific regressors: daily returns, volatility, log of number of trades and a set of day-of-the-week dummy variables, as presented in Table 3. Besides, we include two variables related to information events, both taken from Bloomberg: a dummy for date of earning announcements, and number of newswires. Newswire frequency is measured with a database of number of news releases compiled by Bloomberg in a stock-day basis.

The results in Table 3 indicate that more volatile and actively traded days tend to have less information asymmetry. This is consistent with the direct relation between trading activity and noise trading reported by Berkman and Koch (2008). Besides, the day-of-the week dummies confirm the U-pattern on PIN anticipated in Table 2, where PIN tends to be higher at the start and the end of the week. Most interesting, there is a significant negative relationship with earning announcements, and positive with the log of number of newswires. The first result suggests that earning announcements attract increased noise trading, confirming a finding by EEOW (2008).[7] The second result suggests a novel and intuitive time-series positive relation between news and informed trading activity. The closest previous result, though not directly comparable, is by Sankaraguruswamy et al. (2013) who report a cross-sectional inverse relation between new releases and the static PIN for US firms.

3.4. Testing informed trading effects on liquidity

To investigate for the relation between information and both liquidity and returns we set up panel data models for the whole sample. Doing so increases the ratio between the cross-sectional and time-series dimensions, critical to assure the asymptotic properties of panel data estimation (Driscoll & Kraay, 1998). Country heterogeneity is taken care of by firm-fixed effects in each model. Panel data models that pool together data from different countries have been used in for example by Christoffersen, Chung, and Errunza (2006), who run a panel data of monthly returns for 305 firms coming from 12 emerging markets. Additionally, we

[6] Using the data of Table 1 of Lai et al. (2014, p. 182) we find a sizable and statistically significant difference between the average PIN for developed markets (0.261 for 16.480 firms) and for emerging markets (0.296 for 13.255 firms). Besides there is a strong negative correlation between the reported number of firms and the average PIN by market (−0.42). In particular, the average PINs for Brazil and Mexico (0.288 and 0.317 respectively) are lower than those for Chile, Argentina and Peru (0.318, 0.354 and 0.391, respectively). In turn, Table 1A of Cruces and Kawamura (2005) presents a lower average PIN for Brazil and Mexico (0.16 and 0.17), compared to Argentina, Chile, Colombia and Peru (0.21, 0.22, 0.29, 0.19, respectively) in a cross sectional study on PIN in Latin American markets. Moreover, they report a negative relation between PIN and several country-wide shareholder protection measures.

[7] Specifically, they find that the dynamic PIN increases before earning announcement and decreases afterwards. We find the latter effect but not the former.

Table 2
Summary statistics of the dynamic PIN model and quoted spread by day of the week.

	PIN			Quoted bid-ask spread		
Day	Average	Lower bound	Upper bound	Average	Lower bound	Upper bound
Monday	0.276	0.271	0.281	1.13%	1.10%	1.15%
Tuesday	0.270	0.265	0.275	1.07%	1.05%	1.09%
Wednesday	0.256	0.251	0.260	1.07%	1.04%	1.09%
Thursday	0.261	0.257	0.265	1.09%	1.07%	1.11%
Friday	0.265	0.261	0.270	1.11%	1.09%	1.14%

This table reports the summary statistics of the dynamic Probability of Informed Trading (PIN). PIN is estimated in a stock-day basis as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data from Bloomberg. For each variable the average by day-of-the-week and lower and upper bounds for a 95% confidence interval are reported.

also run panel data models per country to make sure that the results are not country specific. To address further heterogeneity issues we run stock-specific regressions of liquidity on PIN as a robustness test in Section 4.2.

The panel data model is as follows:

$$\log\_liquidity_{it} = \beta_0 + \beta_1 r_{it} + \beta_2 v_{it} + \beta_3 \ln(NT_{it}) + \beta_4 PIN_{it} + \beta_5 PIN_{it-1} + \beta_6 D_{it} + \mu_{it} \quad (6)$$

where $\log\_liquidity_{it}$ is the natural log of one of the two alternative liquidity measures: *quoted spread* (Eq. (4)) and *effective spread* (Eq. (5)), r_{it} is the daily return (close-to-close), v_{it} is the daily volatility as the normalized range of high and low prices during the day, $\ln(NT_{it})$ is the natural log of number of trades, and PIN_{it} is the dynamic PIN measure estimated as explained above. Including the lag of the dynamic PIN allows to test for a persistent effect of the asymmetric information on the effective spread. We also include day-of-the-week effects, D_{it}.

Similar panel data regressions of liquidity on firm characteristics have been used previously by Grullon, Kanatas, and Weston (2004) and Lesmond (2005) and more recently by De Cesari, Espenlaub, and Khurshed (2011) and Hendershott and Moulton (2011). Following that literature, we include returns, volatility and trading activity as control variables, expecting a positive sign for volatility and negative signs for returns and trading activity. Each one of the specifications includes fixed effects and panel-corrected standard errors (PCSEs). Standard errors from PCSE estimation "are robust to very general forms of spatial and temporal dependence as the time dimension becomes large" (Driscoll & Kraay, 1998, p. 549). To validate the panel data regressions, Breusch and Pagan, Hausman, Wooldridge, modified Wald and Breusch and Pagan LM tests were used.

Table 3
Panel data Tobit regression of daily change of dynamic PIN.

Dependent variable:	ΔPIN_{it}
r_{it}	−0.0086
v_{it}	−0.5979***
$\ln(NT_{it})$	−0.0045***
$\ln\_news_{it}$	0.0046***
$Earnings\_Announc_{it}$	−0.0202**
dow_1	0.0177***
dow_2	−0.0055*
dow_4	0.0176***
dow_5	0.0163***
No. of observations	42,695

This table reports the results of a random effects Tobit panel data model to test for the relation between ΔPIN_{it}, the daily change of PIN, on determinants, the measure of dynamic informed trading. PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. i: stock, t: day, r_{it}: close-to-close daily return, v_{it}: daily volatility, NT_{it}: number of trades. $Earnings\_Announc_{it}$ is a dummy variable to be one in stock-days with earning announcements. $\ln\_news_{it} = \ln(1 + news_{it})$ where $news_{it}$ is the number of newswires per stock-day. dow_t: day-of-week dummy, dummy for Wednesday is omitted. The model is estimated for a dependent variable truncated between −1 and 1. *, **, and ***: Statistically significant at the 10%, 5% and 1% levels, respectively.

Table 4
Results for panel data regressions of liquidity measures on PIN.

Dependent variable:	Log of quoted spread	Log of effective spread	Log of effective spread					
Sample:	Full		Argentina	Brazil	Chile	Colombia	Mexico	Peru
r_{it}	−0.485***	−0.698***	1.335***	−0.401***	−1.070***	−1.429***	−1.506***	−0.442***
v_{it}	5.062***	6.898***	8.281***	4.779***	8.416***	16.579***	13.569***	5.742***
PIN_{it}	0.234***	0.285***	0.238***	0.310***	0.291***	0.436***	0.302***	0.152***
PIN_{it-1}	0.063***	0.082***	0.060	0.067***	0.160***	0.046	0.188***	−0.075
$\ln(NT_{it})$	−0.176***	−0.175***	0.251***	−0.142***	−0.213***	0.151***	−0.294***	−0.120***
No. of observations	42,699	42,505	1,512	21,231	7,620	3,002	6,376	2,764
$\overline{R^2}$	0.884	0.859	0.490	0.920	0.651	0.634	0.827	0.477

This table reports the results of the panel data models to test for the relation between liquidity and PIN, the measure of dynamic informed trading as indicated in Eq. (6). PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. i: stock, t: day, r_{it}: close-to-close daily return, v_{it}: daily volatility, NT_{it}: number of trades. Panel data models estimated with fixed effects and using PCSE corrections for auto and cross-correlation and heterocedasticity. All models include day-of-the week variables (omitted). *, **, and ***: Statistically significant at the 10%, 5% and 1% levels, respectively.

3.5. Testing informed trading effects on returns

The dynamic PIN measures the degree of informed trading activity in a market but says nothing about the direction of the information. In response, we define the signed dynamic PIN variable, *Signed_PIN,* taking the dynamic PIN with the sign of the net order imbalance of the day (K in Section 3.1). Thus, if the number of buys is higher (lower) than the number of sells, signed PIN is taken to be positive (negative), inferring a positive (negative) sign on the information of the day. This can be supported on the asymmetric information models of Kyle (1985) and Glosten and Milgrom (1985). A direct precedent is given by Chordia and Subrahmanyam (2004) who theoretically model and empirically report a positive relation between order imbalance and returns. The panel data model is as follows:

$$daily\_return_{it} = \beta_0 + \beta_1 r_{it-1} + index\_return_{jt} + \beta_4 Signed\_PIN_{it} + \beta_5 Signed\_PIN_{it-1} + \beta_6 D_{it} + \mu_{it}. \tag{7}$$

The independent variable, $daily\_return_{it}$, is alternatively the close-to-close daily return r_{it} and the open-to-close daily return r'_{it}. As control variables we include the lagged close-to-close daily return, r_{it-1}, and the return of the main index of each stock market, $index\_return_{jt}$. In some specifications we include the lag of the signed PIN, $Signed\_PIN_{it-1}$, to test for persistence of the effect of information asymmetry on returns. As in the liquidity model above, we run the panel data model for the whole sample and by country, to make sure that the results are not country specific. These panel data models include fixed effects and PCSE corrections for heterocedasticity and cross-correlation in the residuals. We run the usual battery of specification tests to assure proper specification. Finally since the positive effect of PIN on returns is the central result of this study, we subjected it to a battery of robustness tests in Section 4.2 that include time-series stock-specific regressions, instrumental variable regressions, and liquidity adjusted returns.

4. Results

4.1. Informed trading effects on liquidity

Table 4 reports the results for the panel data regressions (Eq. (6)) of daily liquidity measures on informed trading and control variables, with the two alternative liquidity measures for the whole sample and for the effective spread by country. As in previous

Table 5
Results for panel data regressions of daily returns on PIN.

Dependent variable:	r_{it}	r'_{it}	r_{it}	r_{it}					
Sample:	Full	Full	Full	Argentina	Brazil	Chile	Colombia	Mexico	Peru
r_{it-1}:	−0.079***	−0.008	−0.070***	0.113***	−0.123***	0.100***	−0.003	−0.051	0.009
$index\_return_{jt}$	0.564***	0.375***	0.564***	0.605***	0.547***	0.433***	0.751***	0.981***	0.243***
$Signed\_PIN_{it}$	0.011***	0.008***	0.011***	0.015***	0.015***	0.002	0.007***	0.005***	0.014***
$Signed\_PIN_{it-1}$			−0.002***	0.003	−0.002***	−0.002	0.000	−0.004***	−0.004***
No. of observations	43,717	43,717	43,717	1,595	21,604	7,828	3,185	2,833	6,672
$\overline{R^2}$	0.094	0.052	0.094	0.163	0.100	0.038	0.266	0.305	0.101

This table reports the results of the panel data models to test for the relation between daily returns and PIN, the measure of dynamic informed trading, as indicated in Eq. (7). $Signed\_PIN_{it}$ is the PIN_{it} measure multiplied by the sign of the daily order imbalance (buys minus sells). PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. i: stock, t: day, r_{it}: close-to-close daily return, r'_{it}: open-to-close daily return, $index\_return_{jt}$: daily return of the main stock market index. Panel data models estimated with fixed effects and using PCSE corrections for auto- and cross-correlation and heterocedasticity. All models include day-of-the week variables (omitted). *, **, and ***: Statistically significant at the 10%, 5% and 1% levels, respectively.

Table 6
Estimation of the effect of a 1% increase of PIN on daily returns.

Sample	Effect on daily returns, close-to-close			
	One-day effect	Next day reversion	Total two-day effect	p-value
Full	0.0107	−0.0019	0.0089	0.000
Argentina	0.0152	0.0027	0.0180	0.000
Brazil	0.0149	−0.0016	0.0133	0.000
Chile	0.0018	−0.0017	0.0001	0.962
Colombia	0.0072	−0.0002	0.0070	0.000
Mexico	0.0047	−0.0036	0.0010	0.000
Peru	0.0137	−0.0041	0.0096	0.372

This table reports the estimation of the effect of an increase of 1% PIN on daily returns, based on the results presented in Table 5, columns 3 to 9. The one-day (next day) effect corresponds to the coefficient of $Signed\_PIN_{it}$, ($Signed\_PIN_{it-1}$) and the sum of the two, the total two-day effect. $Signed\_PIN_{it}$ is the PIN_{it} measure multiplied by the sign of the daily order imbalance. PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg.

literature, we find a negative effect of returns and trading activity and a positive effect of volatility on the liquidity measures, all of them statistically significant. Thus, Latin American stocks tend to be more liquid in days with positive returns, low volatility and high trading activity. For the most part, these relations appear significant and with the expected sign in the country specific panel data results. Most importantly, each model reports a positive and highly significant effect of information asymmetry (PIN), both on the quoted and effective spreads, consistent with the hypothesis that a higher probability of informed trading should reduce liquidity, as reported on previous empirical studies (Chung et al., 2005; Easley et al., 1997, 1996). This is an out-of-sample verification of the implications of the informed trading models of Kyle (1985) and Glosten and Milgrom (1985), all the most interesting, since none of the Latin America stock markets of the study, except Brazil for the most liquid stocks, have designated market makers as assumed by both informed trading models.

The effect of PIN on liquidity per country, reported in the last six columns, is the lowest for Argentina and Peru, the smallest markets of the sample, and similar for the others. As for economic significance, moving from the 10% to the 90% percentile of the dynamic PIN distribution (see Table 1) leads to an average increase in spreads of 13% in Argentina, a 17% in Brazil, 16% in Chile, 22% in Colombia, 14% in Mexico and 10% in Peru.

The results of Table 4 include a significant positive coefficient of the lagged PIN for the whole sample in both liquidity measures and in three countries. This persistent effect of dynamic PIN is consistent with its reported ability to predict next day liquidity (EEOW; Lei & Wu, 2005).

We also run separated panel data regressions by quartiles of size (market capitalization), trading activity (dollar trading value) and visibility (total number of newswires). Although related, these three measures are not redundant. Across the sample, the Spearman rank correlation between market cap and total news (trading value) is 0.56 (0.77). In unreported regressions, available upon request to the authors, we find a significant effect of PIN on liquidity in each one of the 12 quartiles. Moreover, in panel data regressions for the whole sample that replace the dynamic PIN with interactive variables of dynamic PIN with dummies by quartile, we test for differential effects of PIN on liquidity across quartiles. We only find one statistically different effect: a higher effect of PIN on the liquidity of the lowest size quartile firms.

4.2. Informed trading effects on returns

Table 5 reports the results of panel data regressions of daily returns on informed trading and control variables (Eq. (7)), three for the full sample, six per country. Full sample models show a highly significant positive effect for market returns, which captures, to some extent, the systematic component in the stock return.

Results in Table 5 consistently depict a highly significant positive relation between information asymmetry and both types of daily returns of individual stocks, as predicted by market microstructure theory, with the sole exception of Chile. Economic significance can be estimated with the results of the individual country models of Table 5. A 1% increase on the signed PIN in Argentina or Brazil represents an average additional daily return equivalent to a 3.7% annual return. The corresponding effects are 1.8% for Colombia, 1.2% for Mexico, 3.4% for Peru, and 0.4% for Chile.

As for the persistence of the signed PIN effect on returns, Table 6 shows a statistically significant negative coefficient of the lagged signed PIN for the whole sample and for Brazil, Mexico and Peru, implying some reversion of the PIN effect on prices on the next day. In the other three countries, this effect is not statistically significant. This reversion can be explained from previous empirical studies. PIN, as a proxy of informed trading, is expected to hold a twofold effect on returns: a transient effect due to liquidity restrictions and a permanent effect associated with information (Hasbrouck, 1991a; Duarte & Young, 2009). Following that line of thought, we assume that all the temporal effects of PIN on prices disappear at the end of the next trading day. Consequently, adding up the coefficients of the current and lagged signed PIN in the last seven columns of Table 5, we estimate the permanent effect of the information asymmetry on prices, as summarized in Table 6. Interestingly, the permanent effect of the signed dynamic PIN on returns is highly statistically significant in five out of six countries, which provides support for the dynamic PIN as a proxy for informed trading, not only liquidity effects.

Finally, as in the previous section, we run separated panel data regressions of daily returns by quartiles of size, trading activity and visibility. In unreported regressions, available upon request to the authors, we find a significant effect of signed PIN on daily returns in

Table 7
Robustness tests: Results for stock-specific OLS regressions of liquidity measures on PIN.

	N	r_t		v_t		$\ln(NT_t)$		PIN_t		$\ln(Effective\ bid\_ask\ spread_{t-1})$	
Argentina	12	−1.455	(0.566)**	6.474	(0.874)***	−0.140	(0.028)***	0.251	(0.102)**	0.234	(0.032)***
Brazil	169	−0.305	(0.158)**	7.316	(0.471)***	−0.148	(0.008)***	0.330	(0.077)***	0.270	(0.015)***
Chile	58	−1.394	(0.377)***	12.297	(0.723)***	−0.226	(0.017)***	0.209	(0.055)***	0.229	(0.013)***
Colombia	23	−2.075	(0.562)***	18.334	(1.619)***	−0.164	(0.017)***	0.407	(0.063)***	0.185	(0.024)***
Mexico	50	−1.074	(0.456)**	15.393	(0.945)***	−0.327	(0.022)***	0.316	(0.076)***	0.186	(0.021)***
Peru	21	−0.439	(0.598)	6.206	(1.321)***	−0.134	(0.028)***	0.303	(0.087)***	0.229	(0.032)***
Overall	333	−0.782	(0.140)***	10.057	(0.397)***	−0.188	(0.007)***	0.307	(0.043)***	0.240	(0.009)***

This table reports the mean coefficients of the stock-specific OLS regressions (Eq. (8)) to test for the relation between liquidity, measured by the natural log of the effective bid-ask spread, and PIN, the measure of dynamic informed trading. PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. t: day, r_t: close-to-close daily return, v_t: daily volatility, NT_t: number of trades. N: number of stocks per country. Standard errors of mean coefficients in parenthesis. *, **, and ***: Statistically significant mean coefficients at the 10%, 5% and 1% levels, respectively.

each one of the 12 quartiles. We formally test for differential effects across quartiles estimating panel data regressions for the full sample but replacing signed PIN with interactive variables of signed PIN and dummies by quartile. We find no statistically differential effect between any of the quartiles.

4.3. Robustness tests

To verify the robustness of the relation between liquidity and PIN we estimate the following OLS regression for the daily effective bid-ask spread of each individual stock[8]:

$$\ln(effectives\_pread_t) = \beta_0 + \beta_1 r_t + \beta_2 v_t + \beta_3 \ln(NT_t) + \beta_4 \mathrm{PIN}_t + \ln(effective\ spread_{t-1}). \tag{8}$$

This model addresses two possible criticisms of our panel data results. First, it includes the lagged effective bid-ask spread, since some studies have regarded liquidity as a dynamic variable (e.g. Rhee & Wang, 2009), i.e. past occurrences of liquidity contain information for future values. Second, panel data models, like those in Table 4, assume that coefficients for control variables are invariant across stocks.

The summary of the results of those regressions is presented in Table 7. By large, the mean coefficients of the controls are significant and with the right sign both at each country and in the overall sample. Most importantly, the mean coefficient of PIN is positive and statistically significant at the 5% level for every country and the entire sample. The results confirm a positive relation between informed trading, measured by the dynamic PIN, and liquidity, measured by the effective bid-ask spread.

Similarly, we use the following stock-specific GARCH (1,1) model to test for the relation between closing daily returns and the signed dynamic PIN.

$$\begin{aligned} r_t &= \beta_0 + \beta_1 r_{t-1} + \beta_2 index\_return_{jt} + \beta_3 Signed\_PIN_t + \varepsilon_t \\ \sigma_t^2 &= \theta_0 + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \sigma_{t-1}^2 \end{aligned} \tag{9}$$

This model has several advantages over panel data specification in Eq. (7). First, this approach doesn't assume equal coefficients for the right-hand variables. Second, since we are including the index return, this model captures only idiosyncratic effects of PIN on returns. Third, this model allows for an explicit modeling of the conditional volatility. The mean coefficients for both the mean and variance Eq. (9) grouped by country are presented in Table 8. For the overall sample and four countries, the mean coefficient of the signed PIN is positive and significant at the 5% level, and marginally significant, at the 10%, for the remaining two. These results confirm that informed trading, measured by the signed PIN, moves prices in the direction of the information.

Finally we run two additional robustness tests on the positive relation of signed PIN on daily returns, reported in Table 9. First, there is a valid concern of endogeneity between signed PIN and daily returns. Ideally, we should test whether current signed PIN has any predictive power over next-day returns in the expected direction; however, in any reasonably efficient security market today's information will not significantly predict tomorrow's prices. Instead, we run an instrumental variable regression to address the endogeneity. Lag of signed PIN, the most obvious choice for an instrument, is not a valid instrument because of its significant relation with current daily returns (see Table 5). The chosen instruments are the lag of PIN and the lag of the natural log of volume (in US dollars). We run a two-stage instrumental variable panel data regression of daily returns on the explanatory variables and the "instrumented" signed PIN, the results are reported in the second column of Table 9, along with comparable results of Eq. (7) as reported in Table 8. We still find a significant positive effect of Signed Dynamic PIN on daily returns.[9]

[8] Examples of firm-specific regressions can been found in Kaniel, Ozoguz, and Starks (2012) and Rhee and Wang (2009).

[9] A Hausman test strongly rejects the absence of endogeneity in the original model. An F-Test suggested by Staiger and Stock (1997) with a F statistic = 25.6 validates the mentioned variables as instruments of the signed PIN.

Table 8
Robustness test: Results for stock-specific GARCH (1,1) regressions of liquidity measures on returns.

		Mean equation						Variance equation			
	N	r_{t-1}		$index\_return_{jt}$		$Signed\_PIN_t$		ε^2_{t-1}		σ^2_{t-1}	
Argentina	7	0.025	(0.033)	0.584	(0.081)***	0.014	(0.002) ***	0.571	(0.117) ***	0.094	(0.152)
Brazil	120	−0.049	(0.013)***	0.545	(0.035)***	0.025	(0.005)***	0.290	(0.052)***	0.231	(0.047)***
Chile	49	0.077	(0.017)***	0.402	(0.029)***	0.006	(0.004)*	0.247	(0.053)***	0.347	(0.064)***
Colombia	13	−0.019	(0.029)	0.789	(0.076)***	0.005	(0.002)**	0.272	(0.078)***	0.234	(0.113)**
Mexico	37	−0.046	(0.021)**	0.225	(0.029)***	0.025	(0.005)***	0.208	(0.050)***	0.299	(0.078)***
Peru	18	−0.046	(0.030)*	0.914	(0.052)***	0.004	(0.002)**	0.420	(0.047)***	0.153	(0.050)***
Overall	244	−0.019	(0.009)**	0.509	(0.023)***	0.018	(0.003)***	0.285	(0.030)***	0.255	(0.030)***

This table reports the mean coefficients of the stock-specific GARCH regressions (Eq. (9)) to test for the relation between returns and PIN, the measure of dynamic informed trading. $Signed\_PIN_{it}$ is the PIN_{it} measure multiplied by the sign of the daily order imbalance. PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. t: day, r_t: daily return (close to close), $index\_return_{jt}$: daily return of the main stock market index, ε^2_{t-1} and σ^2_{t-1}, ARCH (1) and GARCH (1) effects in the variance equation. If convergence was not achieved after 50 iterations, the stock regression was discarded. N: number of stocks per country. Standard errors of mean coefficients in parenthesis. *, **, and ***: Statistically significant mean coefficients at the 10%, 5% and 1% levels, respectively.

Second, signed PIN effect on returns might be due to liquidity rather than to information effects. This concern has been partially addressed by the estimation of the two-day effect of signed PIN on returns, as discussed in Section 4.2, which arguably subtracts most of the transient effects. However, for further reassurance we run additional panel data regressions on modified versions of the daily returns adjusted for liquidity effects. Specifically, in the spirit of Roll (1984), we define the adjusted close-to-close daily return, r_adj_{it},using $quoted\_spread_t$, the average quoted bid-ask spread for the day t, as follows:

$$r\_adj_{\mathrm{it}} = \begin{cases} \max\left[0,\quad Ln\left(p_{\mathrm{it}}\ \left(1-\frac{1}{2}quoted\_spread_{it}\right)\Big/\left(p_{it-1}\ \left(1+\frac{1}{2}quoted\_spread_{it-1}\right)\right)\right)\right] & \text{if } r_{it}>0 \\ \min\left[0,\quad Ln\left(p_{\mathrm{it}}\ \left(1+\frac{1}{2}quoted\_spread_{it}\right)\Big/\left(p_{it-1}\ \left(1-\frac{1}{2}quoted\_spread_{it-1}\right)\right)\right)\right] & \text{if } r_{it}<0. \end{cases} \tag{10}$$

This adjustment takes into account that some part of a positive daily close-to-close return is due to the bounce between the yesterday's closing bid and today's closing ask especially if the last trade of today is a buy and the yesterday's last trade was a sell, and the converse for negative daily returns. This conservative adjustment should get rid of any effect on daily returns that might be due to the bid-ask bounce. Likewise, we obtain the adjusted intraday return, rt_{-1}, (open-to close).

We run panel data models similar to Eq. (7) with the adjusted daily and intraday returns, as reported in Table 9. Interestingly, consistent with the intended suppression of the bid-ask spread, the negative effect of lagged return disappears. The effect of signed PIN in the adjusted daily returns is somewhat lower than on daily returns, but still significant and positive, and likewise for the intraday adjusted returns.

5. Conclusions

The econometric models presented here test two basic implications of market microstructure theoretical models of information asymmetry (Kyle, 1985; Glosten & Milgrom, 1985), that information asymmetry should reduce liquidity and move prices in the direction of the information. These implications were tested using the dynamic PIN model of EEOW, which allows for a more rich

Table 9
Robustness test. Alternative panel data regressions of daily returns on PIN.

Dependent variable:	r_{it}	r_{it} IV	r_adj_{it}	r'_{it}	r'_adj_{it}
r_{it-1}:	−0.0699***	−0.0981***	0.0123	−0.0075	0.0321***
$index\_return_{jt}$	0.5640***	0.1519*	0.4111***	0.3747***	0.2866***
$Signed\_PIN_{it}$	0.0107***	0.1239***	0.0069***	0.0084***	0.0065***
$Signed\_PIN_{it-1}$	−0.0019***	−0.0160***	−0.0009***		−0.0010***
No. of observations	43,717	42,676	39,750	43,717	39,576
$\overline{R^2}$	0.0940	0.0263	0.0696	0.0519	0.0414

This table reports the results of the panel data models to test for the relation between returns and PIN, the measure of dynamic informed trading, as indicated in Eq. (7). $Signed\_PIN_{it}$ is the PIN_{it} measure multiplied by the sign of the daily order imbalance (buys minus sells). PIN_{it}, a dynamic measure of informed trading, is estimated as described on EEOW (2008) for a sample of stocks from six Latin American stock markets, in the sample period from August 2, 2010 to March 4, 2011, based on intraday data taken from Bloomberg. i: stock, t: day, r_{it}: close-to-close daily return, r'_{it}: open-to-close daily return, r_adj_{it}: adjusted close-to-close daily return adjusted by bid-ask bounce as indicated in Eq. (10), r'_adj_{it}: adjusted open-to-close daily return adjusted by bid-ask bounce in an analogous way as in Eq. (10), $index\_return_{jt}$: daily return of the main stock market index. Models r_{it} and r'_{it} come from Table 5, and are estimated by panel data regressions with fixed effects and using PCSE corrections for auto and cross-correlation and heterocedasticity. Model r_{it}-IV, is a instrumented variable panel data regression of close-to-close daily returns, estimated by two-stage least-squares within estimator (fixed effects), where $Signed\_PIN_{it}$ is instrumented by the one-day lags of PIN_{it} and the natural log of the trading value in USD. Models r_adj_{it} and r'_adj_{it} are run by a linear panel data regression with fixed effects and heterocedasticity robust errors. All models include day-of-the week variables (omitted). *, **, and ***: Statistically significant at the 10%, 5% and 1% levels, respectively.

dynamic structure of the arrival of informed and uninformed trading than the static PIN models (Easley et al., 1997, 1996). More importantly, the results can be regarded as an out-of-sample test of the dynamic PIN measure as a proxy for informed trading. The positive relationship between the signed dynamic PIN was tested and found statistically significant under a series of different robustness tests that account for stock specific effects, liquidity effects and for the endogeneity between informed trading and returns.

On the other hand, the results portray some diversity across the six stock markets. Argentina's stock market, the smallest by both market cap and trading activity, shows a high sensitivity to asymmetric information effects in returns, in addition to the second highest average PIN and the highest quoted bid-ask spreads. Brazil and Mexico, the two largest markets present the lowest PIN and quoted and effective bid-ask spreads. Chile, an intermediate market by size, reports low sensitivity to asymmetric information in returns. Colombia, another intermediate market, displays high sensitivity of liquidity to returns. Finally, Peru, the second smallest market activity, shows the highest averages of PIN and effective bid-ask spreads, and exhibits low sensitivities of liquidity to asymmetric information. We leave for future research to test whether these differences can be explained by market design, institutions, information availability or different composition of traders.

Acknowledgments

This article originated from a thesis completed by the two last authors in partial fulfillment for the degree of Master Sc. on Finance at Universidad EAFIT under the guidance of the first author. We thank the anonymous referees for their useful suggestions that helped us to improve this paper. We are also grateful to Rafael Bautista, Tao Li, Samuel Mongrut, Carlos Pombo and participants in CLADEA XLVI conference, III World Finance Conference, XII International Finance Conference, and a doctoral seminar in Universidad de los Andes for useful discussions. All errors remain our own responsibility.

References

Akay, O., Cyree, K., Griffiths, M., & Winters, D. (2012). What does PIN identify? Evidence from the T-bill market. *Journal of Financial Markets*, *15*, 29–46.
Aktas, N., de Bodt, E., Declerck, F., & Van Oppens, H. (2007). The PIN anomaly around M&A announcements. *Journal of Financial Markets*, *10*, 169–191.
Barbedo, C.H., Camilo, E., Pereira, R., & Leal, C. (2010). Premium listing segments and information based trading in Brazil. *Academia, Revista Latinoamericana de Administración*, *45*, 1–19.
Barber, B.M., Odean, T., & Zhu, N. (2009). Systematic noise. *Journal of Financial Markets*, *12*, 547–569.
Bekaert, G., Harvey, C., & Lundblad, C. (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, *20*, 1783–1831.
Berkman, H., & Koch, P.D. (2008). Noise trading and the price formation process. *Journal of Empirical Finance*, *15*, 232–250.
Chang, C., & Lin, E. (2015). Cash-futures basis and the impact of market maturity, informed trading, and expiration effects. *International Review of Economics & Finance*, *35*, 197–213.
Chen, Y., & Zhao, H. (2012). Informed trading, information uncertainty, and price momentum. *Journal of Banking and Finance*, *36*, 2095–2109.
Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Market liquidity and trading activity. *Journal of Finance*, *56*, 501–530.
Chordia, T., & Subrahmanyam, A. (2004). Order imbalance, liquidity, and market returns. *Journal of Financial Economics*, *72*, 485–518.
Christoffersen, P., Chung, H., & Errunza, V. (2006). Size matters: The impact of financial liberalization on individual firms. *Journal of International Money and Finance*, *25*, 1296–1318.
Chung, K.H., Elder, J., & Kim, J.C. (2010). Corporate governance and liquidity. *Journal of Financial and Quantitative Analysis*, *45*, 265–291.
Chung, K., Li, M., & McInish, T. (2005). Information-based trading, price impact of trades, and trade autocorrelation. *Journal of Banking and Finance*, *29*, 1645–1669.
Copeland, T. E., & Galai, D. (1983). Information Effects on the Bid-Ask Spread. *Journal of Finance*, *38*, 1457–1469.
Cruces, J., & Kawamura, E. (2005). *Insider trading and corporate governance in America Latina.* Inter-American development bank, research department series*Working Paper No. 3206*.
De Cesari, A., Espenlaub, S., & Khurshed, A. (2011). Stock repurchases and treasury share sales: Do they stabilize price and enhance liquidity? *Journal of Corporate Finance*, *17*, 1558–1579.
Driscoll, J., & Kraay, A. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, *80*, 549–560.
Duarte, J., & Young, L. (2009). Why is PIN priced? *Journal of Financial Economics*, *91*, 119–138.
Easley, D., Engle, R., O'Hara, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics*, *6*, 171–207.
Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *Journal of Finance*, *57*, 2185–2221.
Easley, D., Hvidkjaer, S., & O'Hara, M. (2010). Factoring information into returns. *Journal of Financial and Quantitative Analysis*, *45*, 293–309.
Easley, D., Kiefer, N., & O'Hara, M. (1997). One day in the life of a very common stock. *Review of Financial Studies*, *10*, 805–835.
Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. (1996). Liquidity, information, and infrequently traded stocks. *Journal of Finance*, *51*, 1405–1436.
Easley, D., Lopez de Prado, M.M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies*, *25*, 1457–1493.
Easley, D., & O'Hara, M. (1992). Time and the process of security price adjustment. *Journal of Finance*, *47*, 577–605.
Engle, R., & Neri, B. (2010). The impact of hedging costs on the bid and ask spread in the options market. *Working paper.* New York University.
Glosten, L., & Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*, 71–100.
Goyenko, R., Holden, C., & Trzcinka, C. (2009). Does liquidity measures measure liquidity? *Journal of Financial Economics*, *92*, 153–181.
Grullon, G., Kanatas, G., & Weston, J. (2004). Advertising, breadth of ownership, and liquidity. *Review of Financial Studies*, *17*, 439–461.
Harris, L. (2003). *Trading & exchanges — Market microstructure for practitioners.* Oxford University Press.
Hasbrouck, J. (1991a). Measuring the information content of stock trades. *Journal of Finance*, *46*, 179–207.
Hasbrouck, J. (1991b). The summary informativeness of stock trades: An econometric analysis. *Review of Financial Studies*, *4*, 571–595.
Hendershott, T., & Moulton, P.C. (2011). Automation, speed, and stock market quality: The NYSE's Hybrid. *Journal of Financial Markets*, *14*, 568–604.
Kaniel, R., Ozoguz, A., & Starks, L. (2012). The high volume return premium: Cross-country evidence. *Journal of Financial Economics*, *103*, 255–279.
Kearney, C. (2012). Emerging markets research: Trends, issues, and future direction. *Emerging Markets Review*, *13*, 159–183.
Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, *53*, 1315–1336.
Lai, S., Ng, L., & Zhang, B. (2014). Does PIN affect equity prices around the world? *Journal of Financial Economics*, *114*, 178–195.
Lee, Y., Liu, Y., Roll, R., & Subrahmanyan, A. (2004). Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange. *Journal of Financial and Quantitative Analysis*, *39*, 327–341.
Lee, C., & Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance*, *46*, 733–746.
Lei, Q., & Wu, G. (2005). Time-varying informed and uninformed trading activities. *Journal of Financial Markets*, *8*, 153–181.
Lesmond, D. (2005). Liquidity of emerging markets. *Journal of Financial Economics*, *77*, 411–452.
Lin, E., Lee, C. -F., & Wang, K. (2013). Futures mispricing, order imbalance, and short-selling constraints. *International Review of Economics and Finance*, *25*, 408–423.

Martins, O.S., Paulo, E., & Albuquerque, P.H.M. (2013). Negociação com Informação Privilegiada e Retorno das Ações na BM&F BOVESPA. *Revista de Administração de Empresas*, *53*, 350–362.

Rhee, S.G., & Wang, J. (2009). Foreign institutional ownership and stock market liquidity: Evidence from Indonesia. *Journal of Banking & Finance*, *33*, 1312–1324.

Roll, R. (1984). A simple measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, *39*, 1127–1139.

Sankaraguruswamy, S., Shen, J., & Yamada, T. (2013). The relationship between the frequency of news release and the information asymmetry: The role of uninformed trading. *Journal of Banking and Finance*, *37*, 4134–4143.

Scharfstein, D., & Stein, J. (1990). Herd behavior and investment. *American Economic Review*, *80*, 465–479.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*, 557–586.

Tay, A., Ting, C., Tse, C., & Warachka, M. (2009). Using high-frequency transaction data to estimate the probability of informed trading. *Journal of Financial Econometrics*, *7*, 288–311.

Vega, C. (2006). Stock price reaction to public and private information. *Journal of Financial Economics*, *82*, 103–133.

Villarraga, E., Giraldo, S., & Agudelo, D. (2012). Asimetría en la información y su efecto en los rendimientos en los mercados accionarios latinoamericanos. *Academia, Revista Latinoamericana de Administración*, *50*, 100–117.

World Federation of Exchanges (2010). Market capitalization of affiliated exchanges at the end of 2010. available at: http://www.world-exchanges.org/statistics/annual-query-tool

Yan, Y., & Zhang, S. (2012). An improved estimation method and empirical properties of the probability of informed trading. *Journal of Banking and Finance*, *36*(2), 454–467.