International Review of Financial Analysis 56 (2018) 136–152

Contents lists available at ScienceDirect

International Review of Financial Analysis

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

# The interactions between price discovery, liquidity and algorithmic trading for U.S.-Canadian cross-listed shares☆

Bart Frijns, Ivan Indriawan*, Alireza Tourani-Rad

*Auckland University of Technology Auckland, New Zealand*

ARTICLE INFO

*Keywords:*
Market microstructure
Price discovery
Cross-listings

*JEL classification:*
C32
G15

ABSTRACT

We analyze price discovery dynamics for Canadian companies cross-listed on the NYSE from January 2004 to August 2017. We employ a structural vector autoregression to assess the interactions between price discovery, liquidity and algorithmic trading activity. We observe that over time, the U.S. market is gaining dominance in terms of price discovery. Improvements in liquidity increase a market's contribution to price discovery, and vice versa. We find that algorithmic trading activity is negatively related to price discovery, indicating negative externalities of high-frequency trading. These results are robust to fragmentation in the Canadian financial markets as well as regulatory changes in both the U.S. and Canada.

## 1. Introduction

One of the central functions of financial markets is price discovery, the process by which prices impound new information (Madhavan, 2000). Price discovery is important because it reflects how well a market gathers, interprets, and incorporates new information into prices. It also emphasizes the importance of obtaining the most current information for decision making, i.e. when market participants adjust their expectations on an asset's fundamental value and update their prices. When an asset is listed in multiple markets, price discovery plays an even more important role as information can be incorporated into prices in any market where the security is listed. In such a case, the market which incorporates new information into prices the fastest, has better information processing capacity than other markets and leads in terms of price discovery. Thus, in a multi-market context, price discovery reflects one form of competitiveness of a market relative to others. Such a competitive advantage may attract more investors to that market leading to an improvement in liquidity in that market.

Given the importance of price discovery in a multi-market setting, it is crucial for exchanges to understand which market contributes more to price discovery, and which factors lead to improving a market's contribution to price discovery. Price discovery may shift from one market to another over time for several reasons, one of them being liquidity. Admati and Pfleiderer (1988) explain that a liquid market attracts liquidity traders and that trading will become more concentrated. A liquid market also attracts more informed traders because such a market is "thick" and informed traders can exploit their private information without making large price concessions. At the same time, liquid markets may attract more analysts which further improves the informational environment. Overall, an increase in liquidity could thus lead to an improvement in price discovery for that market. An interesting recent development that can affect price discovery is the upsurge in algorithmic trading (AT). AT accelerates the speed at which traders can detect and exploit price discrepancies among securities, thus it can potentially enhance price discovery. However, such improvement may come at a cost to other traders who are disadvantaged in terms of speed, and may opt to trade in the other market, leading to a reduction in price discovery. These arguments suggest that price discovery will not remain constant, but will vary over time.

Currently, a clear understanding of how price discovery changes over time and what drives such dynamics, is lacking. For instance, the questions of whether price discovery is persistent over time, or whether the dynamics of price discovery is attributable to changes in market liquidity or AT activity are still to be understood. In addition, whether improvements in price discovery lead to more market liquidity is not known. To address these questions, studying price discovery over a

☆ We thank Robert Faff, Yiuman Tse, Robert Webb, the participants at the 2014 Auckland Finance Meeting, the 2017 New Zealand Finance Colloquium, and the seminar participants at Auckland University of Technology, Massey University, Victoria University of Wellington for their helpful comments and suggestions. We thank the Securities Industry Research Centre of Asia-Pacific (SIRCA) for the Best Paper Award during the 2017 New Zealand Finance Colloquium.

* Corresponding author at: Department of Finance, Auckland University of Technology, Private Bag 92006, Auckland 1020, New Zealand
*E-mail address:* ivan.indriawan@aut.ac.nz (I. Indriawan).

https://doi.org/10.1016/j.irfa.2018.01.005
Received 10 April 2017; Received in revised form 22 November 2017; Accepted 8 January 2018
Available online 31 January 2018
1057-5219/ 

longer time period is necessary. Existing studies tend to examine price discovery over relatively short periods.[1] As such, these studies lean towards explaining cross-sectional differences in price discovery, rather than the dynamics of price discovery and liquidity over time. The importance of studying price discovery over longer periods is further emphasized by the changing financial market landscape as a result of, for example, regulatory changes. One such change is the adoption of the Order Protection Rule which was intended to improve fairness in price execution, and to improve the displaying of quotes and access to market data. Such regulation helps create a more integrated market, and may therefore, affect a market's contribution to price discovery.

In this paper, we assess the interactions between price discovery, liquidity and algorithmic trading for a sample of Canadian stocks traded in Canada and the U.S. Our work contributes to the literature in several ways. First, by computing daily measures of the Hasbrouck (1995) information share (IS) and Gonzalo and Granger (1995) permanent-transitory (PT) decomposition over a long period, we are able to explore trends and persistence in price discovery, issues that have hardly been explored in a multi-market context. This also allows us to examine whether the implementation of the Order Protection Rule affected the dynamics of price discovery. Second, we assess how measures of price discovery, liquidity, and AT activity interact with each other over a longer period. Our analyses shed light on what drives price discovery between markets (i.e. whether changes in relative liquidity and AT activity affect the contribution to price discovery of a market), as well as the importance of price discovery for a market (i.e. whether an improvement in price discovery affects liquidity and AT activity).[2] These findings are valuable for exchanges as they indicate what areas they would need to focus on to improve price discovery. Third, from an empirical perspective, we model the interactions between price discovery measures, liquidity, and AT activity using a structural vector autoregression (SVAR). In contrast to the reduced-form Granger causality tests, which measure predictive relationships, the SVAR allows for the identification of contemporaneous interactions among the variables, while at the same time, taking into account the possible endogeneity among them. This is done using the identification through heteroskedasticity approach developed by Rigobon (2003), which was recently implemented by Chaboud, Chiquoine, Hjalmarsson, and Vega (2014).[3]

Applying our model to Canadian stocks listed on the Toronto Stock Exchange (TSX) and cross-listed on the New York Stock Exchange (NYSE) over the period January 2004 to August 2017, we document several important findings. First, we observe that over time, the U.S. market is gaining in terms of price discovery. Second, we find that several measures of liquidity are causally related to price discovery. Improvements in liquidity (an increase in trading volume and a decrease in effective spread) increase an exchange's contribution to price discovery, implying that the market which provides better liquidity will become more important in terms of price discovery. This impact is incorporated instantaneously (within the same day) as well as with a protracted lag (after several days). Conversely, we find that an increase in price discovery leads to improved liquidity, indicating that the market which leads in terms of price discovery becomes more liquid. Third, we find that in the case of cross-listed stocks, algorithmic trading activity negatively affects price discovery. This is in line with the crowding-out effect which has been documented in the literature (Stein, 2009; Gai, Yao, & Ye, 2014; Egginton, Ness, & Ness, 2016). In particular, as high-frequency traders compete aggressively with one another to create latency arbitrage opportunity, they push away other traders who are disadvantaged in terms of speed. Finally, we find that the dynamic relations between price discovery, liquidity and AT activity persist even after we account for the adoption of the Order Proctection Rule both in the U.S. and Canada. Overall, our findings highlight the importance of liquidity for exchanges to improve price discovery, as well as the importance of price discovery to attract more investors. The impact of high-frequency trading on financial markets should be of interest to exchange officials because while it may improve price discovery for the faster traders, the crowding out effect may hinder the price discovery of the market as a whole.

The rest of this paper is structured as follows. Section 2 discusses existing studies on the determinants of price discovery and how our work contributes in this field. In Section 3, we present the data and report descriptive statistics, as well as our measures of liquidity and AT activity. We discuss our measures for price discovery as well as the models for assessing dynamics in price discovery in Section 4. In Section 5, we report our findings. Section 6 concludes.

## 2. Literature review

A market's contribution to price discovery may change over time for various reasons. In this section, we first discuss factors that may contribute to the change in price discovery over time. We then show how these factors can be modeled to assess the dynamics of price discovery in a dual-market scenario.

There is a growing literature examining price discovery of cross-listed stocks. The majority of it focuses on the determinants of price discovery, with liquidity playing an important role. As discussed in Admati and Pfleiderer (1988), one of the motives for trade in financial markets is traders' preference for liquidity. Given that investors have discretion over where and when to trade, they have the tendency to trade in cheaper and more liquid markets, i.e. when the market is “thick” and their trading has little effect on prices. Such a market may attract more traders, leading to information clustering and a shift in price discovery.

One type of liquidity, which is important for price discovery, is trading volume. It is often observed that large trades have a persistent price impact, with trade prices lower after large sales and higher after large purchases. One possible explanation is that increased volume reflects a greater likelihood that demand for a stock comes from informed traders (Stickel & Verrecchia, 1994). Consequently, investors interpret high volume as an indication that the demand underlying a price change is informative, and therefore should get incorporated into prices. Consistent with this view, Hasbrouck (1995) finds a positive and statistically significant relation between the relative trading volume of a sample of 30 Dow stocks and the NYSE's contribution to price discovery. He explains that markets differ in their ability to process information such as that coming from trades. A market which has an informative trading process can shed light on the interpretation of public information, and therefore, leads in terms of price discovery. Similarly, Pascual et al. (2006) find that a market's relative contribution to the price discovery process is related to its trading activity. Using Spanish stocks that are cross-listed on the NYSE, they find that the Spanish Stock Exchange leads in terms of price discovery due to its large trading activity relative to the NYSE as the satellite market.

[1] For instance, Pascual, Pascual-Fuster, and Climent (2006) study Spanish firms cross-listed on the NYSE for the year 2000. Eun and Sabherwal (2003) study Canadian firms cross-listed on the NYSE from February to July 1998, while Chen and Choi (2012) use data from January 1998 to December 2000.

[2] The analysis of the impact of AT activity on price discovery is especially relevant given that AT activity proliferated in the U.S. and Canada at different times. Hence, price discovery between the two markets may have changed over time. In the U.S., high-frequency trading, a subset of AT, became especially popular in 2007 and 2008 (Rogow, 2009, June 19). By 2009, 26 high-frequency traders participate in 68.5% of the dollar volume traded on average (Brogaard, 2010). Gibbs (2007) explains that U.S. players will continue to dominate the market because while Canadian traders ramp up their algorithmic capabilities, they tend to partner with U.S. broker-dealers to leverage their offerings.

[3] The idenfitication through heteroskedasticity approach was recently applied in several finance studies. For example, Chaboud et al. (2014) use the approach to identify the contemporaneous causal impact of AT on triangular arbitrage opportunities. Badshah, Frijns, and Tourani-Rad (2013) use the same approach to assess contemporaneous spillover effects among equity, gold, and exchange rate implied volatility indices. Ehrmann, Fratzscher, and Rigobon (2011) use a similar model to assess international transmission of shocks between money, bond, equity and foreign exchange markets.

Another important determinant of price discovery is the relative bid-ask spread. The trading cost hypothesis predicts that the market with the lower trading costs will react more quickly to new information, as information-based trades are executed where they produce the highest net profit. As a result, a market's contribution to price discovery tends to be inversely related to the bid-ask spread. Consistent with this view, Fleming, Ostdiek, and Whaley (1996) document that index future and option price changes lead price discovery in the stock market because the costs of trading in index futures and options are substantially lower than in the index stocks. Harris, McInish, and Wood (2002) compare the bid-ask spread and a measure of price discovery for the years 1988, 1992, and 1995 for 30 Dow stocks. They find that the NYSE's contribution to price discovery relative to the regional exchanges increases when its spreads relative to the regional markets decline. With regard to the U.S. and Canadian markets, Eun and Sabherwal (2003) explain that the lower spread on U.S. exchanges relative to the TSX represents a competitive threat faced by the TSX liquidity providers from their U.S. counterpart. The TSX liquidity providers who face more competition from U.S. liquidity providers are likely to be more responsive to U.S. prices. Chen and Choi (2012) assess differential private information for Canadian stocks traded in Canada and the U.S. They document that the TSX has more informed trades and a larger information share. This cross-border information imbalance is associated with small but positive price premiums in New York.

In addition to liquidity, studies have looked at how AT activity affects stock markets. Earlier studies document positive aspects of AT on price discovery, particularly on the informativeness of quotes relative to trades. For example, Hendershott, Jones, and Menkveld (2011) assess the impact of quote automation in the NYSE from December 2002 through July 2003. They find that for large stocks in particular, AT enhances the informativeness of quotes by more quickly resetting their quotes after news arrivals, but reduces the trade-related price discovery. Riordan and Storkenmaier (2012) use Deutsche Boerse data from February to June 2007 to study the effect of a latency reduction on price discovery through the introduction of Xetra 8.0 trading platform upgrade. They find that adverse selection has reduced dramatically while the contribution of quotes to price discovery has doubled after the upgrade. Hasbrouck and Saar (2013) use NASDAQ TotalView-ITCH data in the last quarter of 2007 and find that high-frequency trading improves liquidity and price efficiency of the limit order book.

More recent studies, however, highlight some negative aspects of AT on the market as a whole. Stein (2009) explains that recent technological advances resulted in stock markets being dominated by sophisticated professionals using extensive quantitative financial models. Consequently, aggressive investment strategies by these traders have led to a crowding out effect that pushes prices away from their fundamental values, i.e. prices becoming less informative. Gai et al. (2014) explain that since U.S. stock markets impose price, display, and time priority, it is the relative speed, not the absolute speed, that matters. This induces economic incentives not only to invest in speed but also to slow down other traders, which is in line with the “quote stuffing” argument of Egginton et al. (2016), where high-frequency traders submit a profuse number of orders to generate market congestion. Specifically, by submitting large numbers of orders that are canceled very quickly, high-frequency traders may create exploitable latency arbitrage opportunities. Therefore, in contrast to the common notion that competition improves price efficiency, they find that competition through high-frequency trading limits efficiency and inflicts negative congestion externalities on markets as a whole.[4]

[4] Biais, Foucault, and Moinas (2015) find that the improvement in trading speed can either increase or decrease social welfare. In line with this argument, Pagnotta and Philippon (2012) explain that the impact of latency on social welfare depends on the initial level of speed. Particularly, allowing venues to compete on speed improves welfare if the default speed is relatively low, but decreases welfare once the default speed reaches a certain threshold.

It is important to note that in our study, we measure price discovery as a relative term between two different markets. This differs from studies which assess quote and trade informativeness within a single market (e.g. Hendershott et al., 2011; Riordan & Storkenmaier, 2012; Hasbrouck & Saar, 2013) or those which assess the relative contribution of the faster traders relative to other traders within the same market (e.g. Brogaard, Hendershott, & Riordan, 2017). In the case of cross-listed stocks, traders have an option where to execute their orders and make use of their informational advantage. Brogaard et al. (2017) acknowledge this feature of cross-listed stocks when comparing the role of high-frequency and non-high-frequency traders on price discovery in Canada. In their study, they specifically select TSX60 stocks which are not cross-listed in the U.S. because they are unable to precisely measure trading occurring off Canadian exchanges. They note that “it is possible that HFTs' speed and information processing abilities discourage non-HFTs from submitting limit orders” (p. 25). This implies that while price discovery for faster traders may have improved due to high-frequency trading, other traders get pushed away from the market due to speed disadvantage, causing the overall impact of AT on price discovery of a market to be negative.

While there are currently no studies investigating how AT activity affects price discovery across markets, there are a few studies which have looked at how liquidity affects price discovery in a multi-market setting. However, these studies focus predominantly on cross-sectional samples. Harris et al. (2002) study price discovery using a sample of 30 Dow stocks for the years 1988, 1992, and 1995. They calculate differences in price discovery from one year to the next, and relate these differences to changes in the relative spreads between the NYSE and the U.S. regional exchanges. Their findings suggest that higher NYSE spreads reduce the NYSE contribution to price discovery. Frijns, Gilbert, and Tourani-Rad (2010) measure price discovery annually for four Australian stocks cross-listed in New Zealand and five New Zealand stocks cross-listed in Australia from 2002 to 2007. They regress Hasbrouck's (1995) information share on several variables such as the log number of trades in each market, the percentage bid-ask spread in each market, and the log of the market capitalization. They indicate that the growth in the importance of the Australian market is positively related to the growth in the size of the firm and negatively related to the size of the percentage spread in the Australian market. Similarly, Frijns, Gilbert, and Tourani-Rad (2015) measure price discovery annually from 1996 to 2011 for Canadian stocks which are cross-listed on the NYSE, NASDAQ, and AMEX. Their study examines, in particular, the issue of endogeneity between price discovery and measures of liquidity and market quality.

Our work extends the above studies by focusing on the dynamics of price discovery. Specifically, we assess, at a daily frequency, how measures of price discovery, trading volume, bid-ask spread, and AT activity of the U.S. relative to the Canadian markets interact with each other over longer periods. We acknowledge that these variables may be determined simultaneously. For instance, improvements in liquidity and AT activity may lead to a higher contribution to price discovery, while at the same time, higher price discovery may lead to improvements in liquidity and AT activity. To resolve this endogeneity problem, we employ a structural VAR. We follow Chaboud et al. (2014) and account for possible contemporaneous interactions among the VAR variables using the identification through heteroskedasticity approach originally developed by Rigobon (2003).

## 3. Data and descriptive statistics

Our sample consists of Canadian stocks that are traded on the TSX and NYSE from January 2004 through August 2017. Data are collected from the Thomson Reuters Tick History (TRTH) database maintained by Securities Industry Research Centre of Asia-Pacific. We select stocks which are simultaneously traded in both markets through the sample period, had no stock splits, and had a minimum trading history of

3 months preceding the study period. In total, there are 38 stocks which meet these criteria.[5]

We collect intraday data on trade price, trade volume, and the bid and ask quotes at a second and at a millisecond frequency. We use the data at a one-second frequency to compute price discovery measures and construct liquidity measures and use tick-level data sampled at a millisecond frequency to construct the AT proxy.[6] The tick-level data contains all activity observed at the top of the order book. This includes transactions and revisions in bid and ask prices and depths. We omit the first and last 5 min of the trading day to avoid capturing any effects from the open and close of the market. For the U.S. market, we use the national best bid and offer (NBBO) quotes and for the Canadian market, we use quotes posted at the TSX up until January 2011. From 1 February 2011 onwards, we use the Canadian NBBO.[7] Following Grammig, Melvin, and Schlag (2005), we use midpoints of quotes as these are less affected by bid-ask bounce that is normally observed in transaction prices. We also obtain the intraday Canadian-U.S. Dollar exchange rate quotes from TRTH and use the midpoint to convert prices into U.S. dollars. This is to facilitate the specification of the error-term and ensure the comparability of prices between the two markets.

### 3.1. Liquidity measures and algorithmic trading proxy

As measures of liquidity, we use the trading volume and the effective spread. To make inferences about the relations between price discovery and measures of liquidity from both markets, we employ the trading volume and effective spread of the U.S. market relative to the Canadian market. Relative trading volume represents the stock's trading activity and is defined as:

$$Ratio\_Vol_j = \frac{Vol_j^{US}}{Vol_j^{US} + Vol_j^{CAN}}, \tag{1}$$

where $Vol_j^{US}$ and $Vol_j^{CAN}$ are the average U.S. and Canadian trading volume on day $j$, respectively. The second liquidity measure is the relative effective spread, which measures trading costs. Hendershott et al. (2011) explain that effective spreads are more meaningful for the NYSE than quoted spreads because specialists and floor brokers are sometimes willing to trade at prices within the quoted bid and ask prices. The effective spread is measured as:

$$Espread_j = \frac{1}{T}\sum_{t=1}^{T} 2|p_t - m_t|/m_t, \tag{2}$$

where $p_t$ and $m_t$ are the trade price and quote midpoint prevailing at time $t$, respectively. When aggregating over a trading day $j$, we average the effective spreads over $T$ trades. Subsequently, the relative effective spread is computed as:

$$Ratio\_Espread_j = \frac{Espread_j^{US}}{Espread_j^{US} + Espread_j^{CAN}}. \tag{3}$$

For our AT proxy, we use the negative trading volume in hundreds of USD divided by the total message traffic number. This is based on the AT measure of Hendershott et al. (2011) but we implement it as in Boehmer, Fong, and Wu (2015), i.e. by focusing on the top of the limit order book.[8] The proxy can be expressed as:

$$AT_j^i = \frac{-(Dollar\_Vol_j^i)/100}{Total\_messages_j^i}, \tag{4}$$

where $AT_j^i$ is the AT activity for market $i$ on day $j$, $Dollar\_Vol$ is the total dollar trading volume, and $Total\_messages$ is the sum of the number of trade observations and quote changes. This ratio represents the negative dollar volume associated with each trade or quote update. An increase in this measure reflects an increase in algorithmic trading activity. Hendershott et al. (2011) explain that normalization in this case is important because there may be an increase in trading volume over the same interval. Without normalization, a raw message traffic measure may just capture the increase in trading rather than the change in the nature of trading. However, it is important to note that since this AT proxy draws inferences from total message traffic, it makes little distinction between high-frequency traders and slower traders with automated trading systems.[9] Since $AT_j^i$ is negative, relative AT activity is measured as

$$Ratio\_AT_j = \frac{AT_j^{CAN}}{AT_j^{US} + AT_j^{CAN}}. \tag{5}$$

We calculate the liquidity measures and AT proxy for the 38 cross-listed stocks in our sample. Table 1 reports the average market capitalization, trading volume, effective spread, number of messages, and AT activity in both markets, as well as their values in the U.S. relative to Canada.

Our sample covers a broad set of firms with market capitalization ranging from a minimum of \$379 million to a maximum of \$71 billion. On average, daily trading volume in the U.S. is around 1,685,000 shares and in Canada, it is around 1,620,000 shares. This results in a relative trading volume of 51% for the U.S. market, suggesting that trading activity is comparable in both markets. In terms of effective spread, the U.S. market has a lower spread, 9.2 bps compared to 11.1 bps in Canada. Relative effective spread for the U.S. market is 45%, indicating that, on average, trading costs in the U.S. are lower than in Canada. The number of messages per 10-min period differs substantially. In the U.S., there are 1822 messages every 10 min and 3125 messages in Canada, leading to a ratio of 37% for the U.S. market. AT activity, on average, is higher (less negative) in the U.S. compared to Canada with a value of −7.6 and −10.5, respectively. This leads to an AT ratio of 58% for the U.S. relative to Canada.

One of the key variables in our analysis is the AT proxy. To ensure

[5] In comparison, there are only three U.S. cross-listed firms in Canada which fulfill our selection criteria. We also conduct analysis using a more stringent screening by imposing a minimum message count following the approach of Hasbrouck and Saar (2013). A firm is excluded from the sample if more than 10% of the 10-min intervals have fewer than 250 messages (trade and quote). This screening reduces the number of stocks in the sample to 28. As the results are very similar to those discussed in Section 5 and presented in Tables 5–9, we do not report them, but they are available upon request.

[6] Hasbrouck (1995, 2003) indicates that more powerful tests of market efficiency can be carried out by sampling at very high frequencies to reduce the contemporaneous correlation in the reduced form residuals between markets that is created by time aggregation. Hasbrouck (2003) and Hendershott and Jones (2005) use a sampling frequency of 1 s, which produces a low contemporaneous residual correlation and a narrow range of information shares, while Eun and Sabherwal (2003) sample at a 1-min frequency. To assess the sensitivity of our price discovery measures to sampling frequency, we consider both 1 s and 1 min frequency. Using 1 min sampling frequency produces bounds of the information share that are very wide, causing to the IS midpoint to converge to 0.5. The 1 s sampling frequency provides narrow bounds on the information shares, and produces an information share that is generally in line with the PT measure.

[7] With the introduction of the new order protection rule in the Canadian market, the data from 1 February 2011 onwards represents the consolidated tape (this is only available starting early 2011 in Canada, in contrast to the consolidated data being available in the US since the late 1970s). We do note that between 2009 and 2011, the Canadian market fragments considerably. In Appendix B, we test the robustness of our results by taking into account this fragmentation.

[8] The total messages used in our study differ from the one used by Hendershott et al. (2011), who have access to order-level messages. For our sample, we use observations on the exchange's best quotes and trades, rather than all order-related messages. Boehmer et al. (2015) explain that this should not be a problem because the high-frequency trading strategies mentioned in the SEC 2010 concept release involve mostly activity at the top of the book, rather than behind it. They also conducted formal comparison between AT measures based on order-level data used by Hendershott et al. (2011) and AT measures based on best quotes and trades, and find that both measures are highly correlated.

[9] As alternative AT proxies, we use quote-to-trade ratio as well as ratio of limit order duration (see e.g. Hagströmer & Nordén, 2013; Skjeltorp, Sojli, & Tham, 2017). These proxies reflect AT activities as strategies used by algorithmic traders have contributed to a huge increase in the amount of order traffic relative to trade executions, and a decrease in duration between subsequent quote changes. Using either proxies of AT, we find similar results.

**Table 1**

Summary statistics (by firm). This table provides a summary statistics of the 38 stocks in our sample. It reports the company name and the average market capitalization. It also reports the average daily trading volume, the average daily effective spread, the average number of messages per 10-min periods, and the average daily algorithmic trading activity in both markets. Also reported are the ratios of the variables in terms of the U.S. market relative to the total in the U.S. and Canadian markets, e.g. $Ratio\_Vol = Vol^{US} / (Vol^{US} + Vol^{CAN})$.

| | | Mkt Cap | Trading Volume ('000) | | | Effective Spread (bps) | | | Number of Messages (10min) | | | Algorithmic Trading Activity | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | Company | (USD mil) | US | CAN | RATIO | US | CAN | RATIO | US | CAN | RATIO | US | CAN | RATIO |
| 1 | Barrick Gold | 26,601 | 8944 | 3355 | 70% | 4.9 | 5.8 | 46% | 6331 | 9232 | 45% | −14.9 | −7.9 | 28% |
| 2 | Agnico-Eagle Mines Limited | 6682 | 2008 | 877 | 69% | 4.8 | 6.7 | 43% | 2737 | 4737 | 42% | −10.2 | −3.7 | 25% |
| 3 | Agrium Inc. | 9698 | 1368 | 708 | 60% | 4.7 | 6.5 | 44% | 1691 | 2733 | 40% | −13.1 | −5.7 | 29% |
| 4 | BCE Inc. | 28,677 | 634 | 2444 | 23% | 3.9 | 4.3 | 48% | 1425 | 3174 | 37% | −5.1 | −16.9 | 64% |
| 5 | Bank of Montreal | 33,045 | 400 | 1605 | 18% | 3.9 | 3.6 | 51% | 1525 | 3440 | 39% | −3.7 | −23.3 | 72% |
| 6 | Bank of Nova Scotia | 52,240 | 388 | 2170 | 14% | 4.7 | 3.8 | 53% | 1624 | 4189 | 36% | −3.1 | −24.7 | 76% |
| 7 | Brookfield Office | 7815 | 1551 | 540 | 74% | 7.1 | 9.9 | 43% | 1142 | 1484 | 46% | −7.0 | −1.9 | 22% |
| 8 | Cameco Corp. | 8720 | 1889 | 1332 | 57% | 6.2 | 7.4 | 46% | 1922 | 3242 | 41% | −10.9 | −8.3 | 35% |
| 9 | Canadian Imperial Bank Communication | 27,364 | 250 | 1278 | 14% | 4.1 | 3.7 | 52% | 1134 | 2402 | 39% | −3.6 | −26.8 | 77% |
| 10 | Canadian National Railway Company | 32,942 | 919 | 1094 | 44% | 3.0 | 4.0 | 44% | 1722 | 3056 | 41% | −11.1 | −11.0 | 46% |
| 11 | Canadian Natural Resources Ltd. | 31,947 | 2320 | 2575 | 46% | 4.0 | 4.9 | 46% | 3905 | 6544 | 42% | −11.0 | −13.7 | 46% |
| 12 | Canadian Pacific | 14,086 | 612 | 590 | 49% | 4.7 | 6.0 | 45% | 1042 | 1624 | 42% | −13.4 | −10.3 | 42% |
| 13 | Encana Corp. | 24,901 | 4768 | 3235 | 56% | 5.6 | 6.5 | 45% | 3421 | 5118 | 44% | −14.0 | −15.0 | 39% |
| 14 | Enbridge Inc. | 24,319 | 618 | 1299 | 27% | 4.9 | 5.3 | 48% | 1353 | 3014 | 40% | −4.2 | −9.9 | 61% |
| 15 | Enerplus Corp. | 3902 | 755 | 684 | 57% | 9.0 | 10.6 | 46% | 934 | 1 861 | 39% | −9.8 | −4.3 | 34% |
| 16 | Goldcorp Inc. | 19,819 | 6273 | 2970 | 67% | 4.9 | 6.0 | 46% | 5357 | 8479 | 45% | −12.4 | −7.2 | 31% |
| 17 | Gildan Activewear Inc. | 4032 | 440 | 456 | 45% | 6.8 | 8.4 | 46% | 686 | 1219 | 40% | −5.2 | −4.4 | 43% |
| 18 | Kinross Gold Corp. | 7783 | 6569 | 4529 | 57% | 16.5 | 17.5 | 48% | 2750 | 3458 | 47% | −6.5 | −6.7 | 41% |
| 19 | Manulife Financial Corp. | 35,455 | 1610 | 4079 | 28% | 5.7 | 6.4 | 47% | 2487 | 4075 | 42% | −7.1 | −16.1 | 64% |
| 20 | Nexen Inc. | 11,896 | 2151 | 1830 | 45% | 6.2 | 7.3 | 46% | 1989 | 2 615 | 46% | −9.0 | −15.2 | 50% |
| 21 | Potash Corporation of Saskatchewan Inc. | 25,764 | 4808 | 1498 | 73% | 3.9 | 4.9 | 45% | 4142 | 5919 | 43% | −23.8 | −6.4 | 22% |
| 22 | Rogers Communication Inc. | 14,817 | 313 | 1346 | 19% | 4.1 | 5.8 | 45% | 1079 | 1923 | 40% | −7.7 | −14.9 | 65% |
| 23 | Royal Bank of Canada | 71,359 | 650 | 2663 | 17% | 3.4 | 3.4 | 49% | 2113 | 4966 | 38% | −4.7 | −26.6 | 72% |
| 24 | Sun Life Financial | 19,481 | 345 | 1413 | 19% | 5.0 | 5.5 | 48% | 1225 | 2699 | 38% | −3.4 | −13.1 | 70% |
| 25 | Suncor Energy Incorporated | 41,802 | 4080 | 3408 | 53% | 3.7 | 4.5 | 46% | 4944 | 7882 | 43% | −16.0 | −12.3 | 41% |
| 26 | Toronto-Dominion Bank | 60,030 | 755 | 2181 | 22% | 3.4 | 3.4 | 50% | 2210 | 4992 | 39% | −5.5 | −23.0 | 68% |
| 27 | Talisman Energy Inc. | 13,617 | 3501 | 3550 | 47% | 7.4 | 8.6 | 46% | 2308 | 3124 | 45% | −8.5 | −11.6 | 47% |
| 28 | TransCanada Corp. | 24,407 | 442 | 1458 | 22% | 4.4 | 4.6 | 49% | 1243 | 2649 | 38% | −4.2 | −12.7 | 66% |
| 29 | Celestica Inc. | 1769 | 767 | 740 | 49% | 11.5 | 13.5 | 46% | 572 | 939 | 42% | −4.5 | −3.0 | 40% |
| 30 | COTT Corp. | 988 | 496 | 263 | 66% | 18.6 | 23.8 | 45% | 382 | 539 | 44% | −4.1 | −1.7 | 27% |
| 31 | CGI Group | 5857 | 145 | 942 | 15% | 11.4 | 11.1 | 51% | 575 | 1046 | 42% | −1.5 | −6.3 | 76% |
| 32 | Kingsway Financial Services Inc. | 379 | 34 | 96 | 45% | 74.6 | 99.8 | 47% | 76 | 108 | 47% | −0.7 | −3.0 | 51% |
| 33 | MI Developments Inc. | 1385 | 111 | 49 | 74% | 14.0 | 21.8 | 40% | 130 | 182 | 43% | −5.6 | −1.1 | 19% |
| 34 | Precision Drilling Trust | 2612 | 1406 | 1413 | 52% | 12.8 | 14.4 | 46% | 1104 | 1929 | 42% | −7.4 | −5.8 | 39% |
| 35 | Pengrowth Energy Corp. | 2551 | 1110 | 1065 | 57% | 23.4 | 28.8 | 45% | 520 | 983 | 45% | −9.1 | −7.7 | 39% |
| 36 | Ritchie Brothers Auctioneers | 2296 | 331 | 113 | 77% | 8.2 | 14.3 | 41% | 468 | 949 | 41% | −5.7 | −1.3 | 17% |
| 37 | Shaw Communications Inc. | 7982 | 215 | 884 | 20% | 7.0 | 8.3 | 47% | 697 | 1405 | 39% | −2.4 | −6.3 | 68% |
| 38 | TransAlta Corp. | 3651 | 57 | 835 | 7% | 16.1 | 11.5 | 55% | 286 | 817 | 37% | −0.7 | −9.3 | 85% |
| | Mean | | 1685 | 1620 | 51% | 9.2 | 11.1 | 45% | 1822 | 3125 | 37% | −7.6 | −10.5 | 58% |

that this proxy captures algorithmic trading activity, we compare the distribution of message traffic in January 2004 and August 2017. Table 2 reports statistics on the message traffics in the U.S. and Canada. It reports the daily averages for the number of messages, the number of quote observations, the number of trades, the quote-to-trade ratio, and the limit order duration (in seconds), each with a percentage change over the sub-periods. We make several important observations. Referring first to Panel A and B, we observe that by 2017, total messages have grown exponentially by 2963 % in the U.S. and by 5860 % in Canada. This growth is consistent with high-frequency traders playing an increasingly important role in both markets. Most of the growth in total messages comes from quotes rather than trades, particularly in Canada where the number of quotes have grown by 8207 % by 2017, as compared to 3234 % in the U.S. This is reflected in the increase in quote-to-trade ratio over the two sample periods by 179% in the U.S., and by 3030 % in Canada. Limit order duration has also decreased substantially in both markets.

Panel C shows the limit order activity in the U.S. relative to Canada. The second to the last row of Panel C in Table 2 indicates that in 2004, QT ratio in the U.S. is 61% (or 39% for Canada). By 2017, this number has changed to 41% in the U.S. (or 59% in Canada). This figure suggests that over time in Canada, there are more quotes submitted to the limit order book per each trade executed. Furthermore, the last row of Panel C refers to the limit order duration (time taken to update quotes, in seconds) of the U.S. relative to Canada. In 2004, the limit order duration of the U.S. relative to Canada is 43%. This means it takes shorter time for quotes to update in the U.S. than in Canada. By 2017, this figure has changed to 65%, suggesting that it now takes longer time for quotes to update in the U.S. than in Canada. Based on these observations, we conclude that the revision in quotes is more frequent than trades, which is consistent with Hasbrouck and Saar (2013) who associate high-frequency traders with fast order submissions and cancellations.

Fig. 1 plots the 20-day moving average of trading volume, effective spread, and AT activity of the U.S., Canada and their relative values. Panel A shows that relative trading volume, *Ratio_Vol*, has an upward trend. The increase is notable from 2004 to 2008 prior to the Global Financial Crisis when U.S. trading volume peaked. Relative trading volume declined in 2011 before it increased again from 2014 onwards.

Panel B plots the relative effective spread, *Ratio_Espread* over the

**Table 2**
Descriptive statistics of the number of quote-change and trade messages. This table reports the percentage change in daily total messages, number of quote changes, number of trades, quote-to-trade ratio, and limit order duration (in seconds) from January 2004 to August 2017. The numbers reported are averages across the 38 stocks in our sample. Panel A reports the statistics for the U.S. market. Panel B reports the statistics for the Canadian market. Panel C reports the U.S. statistics relative to the total from both markets, e.g. $QT^{US}/(QT^{US}+QT^{CAN})$. *** denotes significance at the 1% level.

| | Jan-04 | Aug-17 | %Change | t-stat |
|---|---|---|---|---|
| *Panel A: US* | | | | |
| Total messages | 3468 | 105,691 | 2963%*** | (9.89) |
| Quote | 2822 | 97,661 | 3234%*** | (10.22) |
| Trade | 646 | 8030 | 2942%*** | (2.67) |
| Quote-to-trade ratio | 4.4 | 18.1 | 179%*** | (4.93) |
| Limit order duration | 11.5 | 1.2 | −92%*** | (−38.07) |
| *Panel B: CAN* | | | | |
| Total messages | 3649 | 190,207 | 5860%*** | (6.91) |
| Quote | 2688 | 181,168 | 8207%*** | (6.35) |
| Trade | 961 | 9038 | 1082%*** | (6.44) |
| Quote-to-trade ratio | 2.8 | 48.3 | 3030%*** | (6.30) |
| Limit order duration | 15.0 | 0.8 | −97%*** | (−34.73) |
| *Panel C: US/(US + CAN)* | | | | |
| Total messages | 0.49 | 0.36 | −25%*** | (−7.57) |
| Quote | 0.51 | 0.35 | −31%*** | (−8.73) |
| Trade | 0.40 | 0.45 | 79%*** | (3.36) |
| Quote-to-trade ratio | 0.61 | 0.41 | −31%*** | (−6.74) |
| Limit order duration | 0.43 | 0.65 | 49%*** | (5.90) |

years. For most of our sample period, the relative effective spread is lower than 0.50, suggesting that trading costs in the U.S. are generally lower than in Canada. Effective spreads in both markets have converged over time, particularly between 2008 and 2011. The plots for the respective markets further show that spreads in the U.S. are consistently lower than in Canada.

Panel C plots AT activity of the U.S. relative to Canada. The plot for the *Ratio_AT* shows a downward sloping trend over the years. This can be attributed to the Canadian market increasing its AT activity over the recent years, especially after the emergence of alternative trading systems in mid-2007 to compete with the TSX (Clark, 2011). Where the U.S. used to report higher AT activity than Canada (ratio of greater than 0.5) before 2008, it has declined to the point where Canadian AT activity is now higher relative to the U.S.

## 4. Methodology

### 4.1. Measuring price discovery

The study of price discovery relies on the assumption that when a security is listed in multiple markets, prices in these markets share a common trend, i.e., prices are cointegrated. Cointegration implies that prices can deviate from each other in the short-run due to frictions, but are bound together in the long-run. In our dual-market case, such a relation can be presented by two $I(1)$ price series, $y_t^{US}$ and $y_t^{CAN}$ being cointegrated with a cointegrating vector $\beta' = (1\ -1)$. The Engle-Granger Representation Theorem states that a cointegrated system can be expressed as an error-correction model. Hence, the stationary process, $\beta' y_t = y_t^{US} - y_t^{CAN}$, can be applied as an error-correction term for the following Vector Error-Correction Model (VECM),

$$\Delta y_t = c + \alpha\beta' y_{t-1} + \sum_{n=1}^{N} \Gamma_n \Delta y_{t-1} + \epsilon_t. \tag{6}$$

where $\Delta y_t$ is the $(2 \times 1)$ vector of log returns, $c$ is a vector of constants, $\alpha$ is a $(2 \times 1)$ vector that measures the speed of adjustment to the error-correction term (i.e. $\alpha' = (\alpha^{US}\ \alpha^{CAN})$), $\Gamma_n$ are $(2 \times 2)$ matrices of AR coefficients, and $\epsilon_t$ is a $(2 \times 1)$ vector of innovations. The VECM has two parts: the first part, $\beta' y_{t-1}$, represents the long-run equilibrium between the price series. The second part, $\sum_{n=1}^{N} \Gamma_n \Delta y_{t-1}$, represents the short-term dynamics induced by market imperfections.

We use the above VECM to compute the price discovery measures between two markets. Our price discovery measures are the Gonzalo and Granger (1995) permanent–transitory (PT) decomposition, and the Hasbrouck (1995) information share (IS). Both are directly related and both measures are derived from the VECM.

The PT measure is concerned with permanent shocks that result in a disequilibrium as markets process news at different speeds. It measures each market's contribution to the common factor, where the contribution is a function of the speed of adjustment coefficients, $\alpha$. Hence, the PT can be computed using the following equation,

$$PT^{US} = \frac{\alpha^{CAN}}{(\alpha^{CAN} + |\alpha^{US}|)}, \tag{7}$$

where $\alpha^{US}$ is negative, and $\alpha^{CAN}$ is positive given our definition of $\beta' = (1\ -1)$. This ratio provides an indication of the degree of dominance of one market over the other market. A higher value of this ratio reflects a greater feedback or contribution from the U.S. Therefore, a $PT^{US}$ of zero implies that the NYSE does not contribute to the price discovery of the stocks, whereas a $PT^{US}$ greater than zero implies feedback from the NYSE to the TSX. $PT^{CAN}$ can be computed as $1 - PT^{US}$.

The IS measures the proportion of variance contributed by one market with respect to the variance of the innovations in the common efficient price. To assess this, note that we can rewrite Eq. (6) as a vector moving average (Wold representation):

$$\Delta y_t = \Psi(L)e_t, \tag{8}$$

where $\Psi(L)$ is a matrix polynomial in the lag operator ($\Psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \cdots$ ). Following the Beveridge and Nelson (1981) decomposition, which states that every (matrix) polynomial has permanent and transitory structure, we can write Eq. (8) in its integrated form as:

$$y_t = \Psi(1) \sum\nolimits_{s=1}^{t} e_s + \Psi^*(L)e_t. \tag{9}$$

where $\Psi(1)$ is the sum of all moving average coefficients, and measures the long-run impact of an innovation to the level of prices. Since prices are cointegrated, $\beta' y_t$ is a stationary process, which implies that $\beta'\Psi(1) = 0$, i.e. the long-run impact is the same for all prices. If we denote $\psi = (\psi^{US}\ \psi^{CAN})$ as the common row vector in $\Psi(1)$, Eq. (9) becomes:

$$y_t = \iota\psi \sum\nolimits_{s=1}^{t} e_s + \Psi^*(L)e_t. \tag{10}$$

The increment $\psi e_t$ in Eq. (10) is the component of the price change that is permanently impounded into the price and is due to new information. Hasbrouck (1995) decomposes the variance of the common factor innovations, i.e., $var(\psi e_t) = \psi\Omega\psi'$. The information share of a market is defined as the proportion of variance in the common factor that is attributable to innovations in that market. Since Hasbrouck (1995) uses the Cholesky factorization of $\Omega = MM'$ to handle contemporaneous correlations, where $M$ is a lower triangular matrix, the information share of market $i$ is defined as:

$$S_i = \frac{([\psi M]_i)^2}{(\psi\Omega\psi')}. \tag{11}$$

The Cholesky decomposition of $\Omega$ orthogonalizes the innovation terms and assigns all common variance to one market. To account for multiple markets, Hasbrouck (1995) suggests that different orderings of the innovation terms be used so that upper and lower information share bounds can be computed. Specifically, we reverse the order of the $\Psi(1)$ as well as $M$ and recompute Eq. (11). The midpoint of these bounds is the IS value.

### 4.2. Modeling price discovery dynamics

Section 2 indicates that factors such as trading volume, bid-ask spread, and algorithmic trading activity may be related to price discovery. If such relations exist, the ratio of those variables in one market

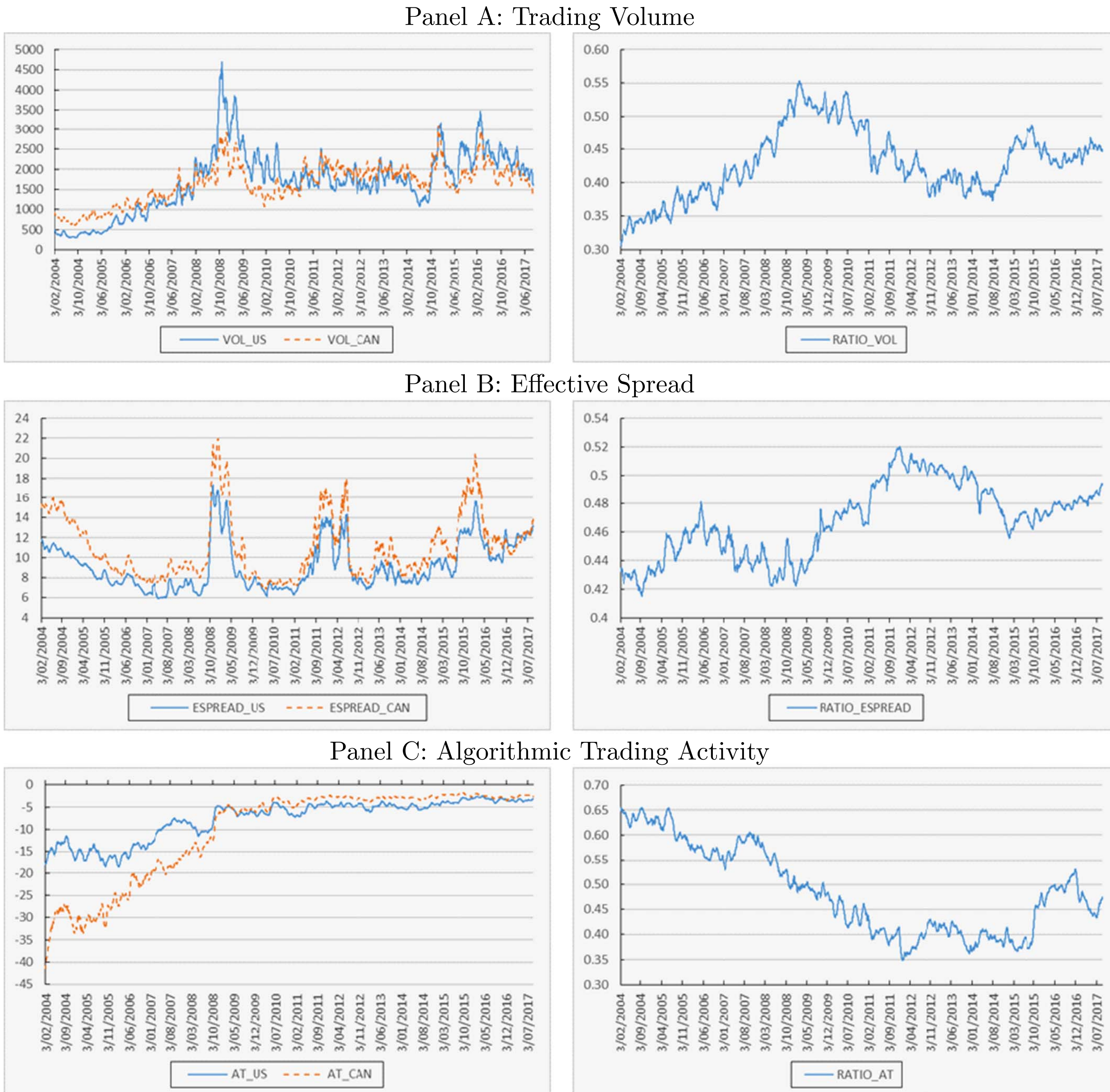


**Fig. 1.** Trading Volume, Effective Spread, and Algorithmic Trading Activity Over Time. This Figure shows time series plots of the U.S. relative daily trading volume, U.S. relative daily effective spread, and U.S. AT activity. The figures are the 20-day moving averages computed from the mean *Ratio_Vol*, *Ratio_Espread*, and *Ratio_AT* for the 38 firms in the sample, respectively. The x-axis represents the sample period from January 2004 to August 2017, while the y-axis represents the value of the levels for each respective variable.

relative to another may affect the dynamics of price discovery between the two markets. To examine such dynamics, we use a VAR to model the interactions between price discovery measures, trading volume, bid-ask spread, and AT activity. We estimate both a reduced-form and a structural VAR (SVAR) that uses the identification through heteroskedasticity approach developed by Rigobon (2003). Doing so, we are able to assess lagged and contemporaneous interactions among the VAR variables. In this section, we focus on the mechanics of the identification through heteroskedasticity methodology. Further details on the methodology are provided in Appendix A.

Given that price discovery measures, trading volume, bid-ask spread, and AT activity may have contemporaneous effects on each other, and assuming these variables exhibit persistence, the dynamics of price discovery can be expressed by the following SVAR:

$$A\Delta Y_t = c + \sum_{k=1}^{K} \Pi_k \Delta Y_{t-k} + \varepsilon_t, \tag{12}$$

where $\Delta Y_t$ is the $(4 \times 1)$ vector of changes in variables, i.e. $\Delta Y_t = (\Delta IS_t, \Delta Ratio\_Vol_t, \Delta Ratio\_Espread_t, \Delta Ratio\_AT_t)'$, $\Pi_k$ is a $(4 \times 4)$ matrix of coefficients for the autoregressive terms for lag $k$, and $\varepsilon_t$ is a vector of error terms. Matrix $A$ captures the structural parameters and is normalized such that all diagonal elements are equal to 1, and its off-diagonal elements capture the contemporaneous interactions between the variables, i.e.,

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} \\ a_{21} & 1 & a_{23} & a_{24} \\ a_{31} & a_{32} & 1 & a_{34} \\ a_{41} & a_{42} & a_{43} & 1 \end{pmatrix}.$$

The off-diagonal elements capture the interactions among the variables. For instance, $a_{12}$, $a_{13}$, $a_{14}$ represent the contemporaneous impact of $\Delta Ratio\_Vol$, $\Delta Ratio\_Espread$ and $\Delta Ratio\_AT$ on $\Delta IS$, while $a_{21}$, $a_{31}$, $a_{41}$ represent the contemporaneous impact of $\Delta IS$ on $\Delta Ratio\_Vol$, $\Delta Ratio\_Espread$ and $\Delta Ratio\_AT$.

Since the contemporaneous relations among the VAR variables are not equal, $A$ is not symmetric. Consequently, the parameters in $A$ cannot be obtained using OLS. To overcome this issue, we estimate Eq. (12) using the identification through heteroskedasticity methodology. This approach starts with transforming Eq. (12) into its reduced-form below:

$$\Delta Y_t \ = A^{-1}c + A^{-1}\sum_{k=1}^{K}\Pi_k \cdot \Delta Y_{t-k} + A^{-1}\varepsilon_t \Delta Y_t \ = \tilde{c} + \sum_{k=1}^{K}\widetilde{\Pi}_k \cdot \Delta Y_{t-k} + \tilde{\varepsilon}_t, \quad (13)$$

where the residuals $\tilde{\varepsilon}_t$ from the reduced-form VAR are related to the residuals $\varepsilon_t$ from the SVAR through the inverse of $A$. Here, matrix $\widetilde{\Pi}_k$ allows us to test for Granger causality among the VAR variables. Since Eq. (13) can be estimated by OLS, it serves as the basis for the heteroskedasticity identification scheme. In particular, the residuals from Eq. (13) can be used to identify different variance regimes, i.e. split $\tilde{\varepsilon}_t$ into different subsamples, such that the covariance matrices under these subsamples are not proportional to each other.[10] Once the different heteroskedastic regimes have been identified, we can increase the number of available moment conditions and use them to estimate the parameters in $A$. The empirical execution of the heteroskedasticity identification scheme can be found in Appendix A.

## 5. Empirical findings

In this section, we begin by showing how price discovery measures for Canadian cross-listed stocks vary over time. We then present the Granger causality results from the reduced-form VAR and the results from the structural VAR as formal approaches to assess the dynamics of price discovery. Finally, we examine whether the adoption of the Reg. NMS affected the dynamics of price discovery between the U.S. and Canadian markets.

### 5.1. *Price discovery over time*

To obtain price discovery estimates over time, the IS and PT are estimated daily for each firm.[11] The daily estimation eliminates the overnight price jumps which typically generate excessive noise. Throughout this paper, our price discovery estimates are based on the U.S. portion of IS and PT. The VECM in Eq. (6) is estimated by applying OLS with optimal lag length suggested by the Schwartz Information Criterion.

Table 3 reports the descriptive statistics of the PT and IS. Panel A reports the statistics for the levels. During the entire sample, the average (median) IS for the U.S. market is 57.9% (62.6%), while for PT, it is 63.7% (66.3%). These figures indicate that the U.S. contribution to price discovery tends to be higher than the Canadian contribution. We observe a wide range in price discovery measures, from 29.6% to 73.5%, and from 39.0% to 78.7% at the 5*th* and 95*th* percentile for IS and PT, respectively. Both measures are negatively skewed, but do not display excess kurtosis. The autocorrelation (AC) for IS and PT are 0.805 and 0.797 for the first lags, and decrease with increasing lags, hence indicating autoregressive processes. The Augmented Dickey Fuller (ADF) test statistics are insignificant, suggesting that unit roots are present in the IS and PT series.

**Table 3**
Descriptive Statistics of the price discovery measures. This table reports the descriptive statistics for the price discovery measures. *IS* and *PT* are estimated daily from January 2004 to August 2017. The figures reported are the averages for all 38 Canadian cross-listed stocks in the sample. Panel A reports statistics for the levels, and Panel B reports statistics for the first differences. ADF is the t-statistics for the Augmented Dickey-Fuller test. *** denotes significance at the 1% level.

| | IS | PT |
|---|---|---|
| *Panel A: Summary Statistics for levels* | | |
| Mean | 0.579 | 0.637 |
| 5th | 0.296 | 0.390 |
| Median | 0.626 | 0.663 |
| 95th | 0.735 | 0.787 |
| Std. dev. | 0.205 | 0.185 |
| Skewness | −0.836 | −0.634 |
| Kurtosis | 3.012 | 2.948 |
| AC | 0.805 | 0.797 |
| ADF | −2.140 | −2.181 |
| *Panel B: Summary Statistics for 1st difference* | | |
| Mean | 0.000 | 0.000 |
| 5th | −0.077 | −0.080 |
| Median | 0.000 | 0.000 |
| 95th | 0.081 | 0.081 |
| Std. dev. | 0.117 | 0.109 |
| Skewness | −0.013 | −0.013 |
| Kurtosis | 7.548 | 7.021 |
| AC | −0.454 | −0.448 |
| ADF | −61.690*** | −60.913*** |

Panel B reports summary statistics for the first differences. The mean values of the first differences are close to zero, although there is quite some variation on a daily basis as can be seen from the range of the 5*th* and 95*th* percentile and the standard deviation. The series have skewness values close to zero with excess kurtosis, suggesting that observations occur predominantly around the mean. We do not observe the first differences to be serially correlated as the AC quickly drops to zero after one lag. Furthermore, the ADF test statistics are highly significant, confirming that the first difference series for IS and PT are stationary.

In Fig. 2, we plot IS and PT from January 2004 to August 2017, based on a 20-day moving average for the 38 stocks in our sample. The IS and PT track each other closely with the PT being consistently higher than the IS. According to both measures, price discovery for the U.S. is lower than 50% prior to 2007. This is consistent with earlier studies which show that the home market for the Canadian-U.S. cross-listed stocks dominate in terms of price discovery.[12] We observe a sharp increase in price discovery around the year 2007. From 2007 onwards, the U.S. market gains dominance with IS and PT greater than 50%. The IS and PT reach around 80% in 2010. One possible explanation for the increase in the U.S.'s contribution to price discovery is the implementation of the Order Protection Rule as part of the Reg. NMS which started in 2006 and was finalized in October 2007, an explanation we examine in Section 5.5.

Apart from the slight decrease in IS and PT in late 2008, the increasing trend in price discovery measures does not seem to be substantially affected by the Global Financial Crisis. Overall, Fig. 2 illustrates that price discovery, as measured by IS and PT, exhibits

[10] Rigobon (2003) suggests that at least two distinct variance regimes for the error terms are required in order for the identification scheme to work.

[11] Prior to estimating the IS and PT, we conduct the usual procedures of unit root and cointegration tests. First, we perform non-stationarity tests using the Augmented-Dickey Fuller test using SIC to select optimal lag length. For all stocks, we cannot reject the presence of a unit root. Subsequently, we conduct Johansen's (1988) test for cointegration. In all tests, we reject the null of no cointegration in favor of the alternative of one cointegrating vector. Since the price series in our sample satisfy both conditions, we conclude that each pair of our sample stocks is cointegrated.

[12] See for example, Eun and Sabherwal (2003) Chen and Choi (2012).

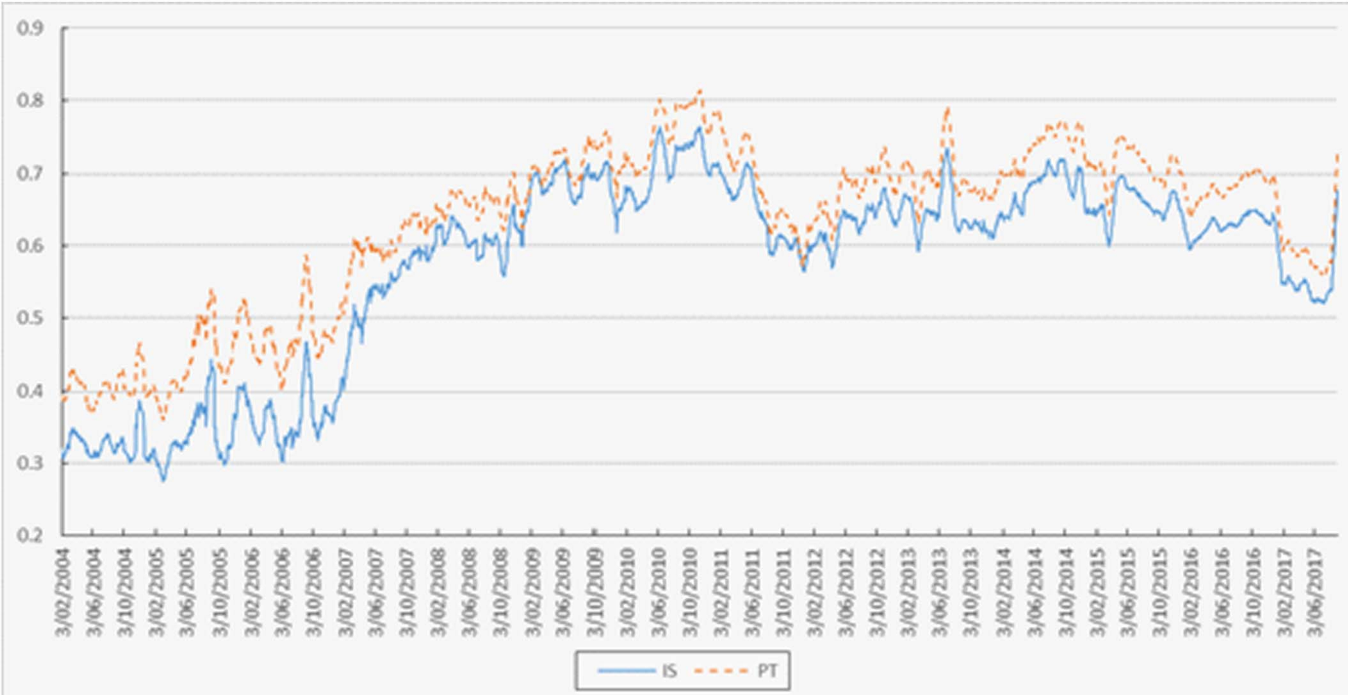


**Fig. 2.** Price discovery measures over time. This Figure shows time series plots of the IS and PT for the U.S. market over the sample period January 2004 to August 2017. The figures are the 20-day moving averages computed from the mean IS and PT for the 38 firms in the sample.

**Table 4**
Correlation matrix between VAR variables. This table presents the correlation matrix for the series *ΔIS*, *ΔPT*, *ΔRatio_Vol*, *ΔRatio_Espread*, and *ΔRatio_AT*. *ΔIS* and *ΔPT* are the first differences in the price discovery measures *IS* and *PT*, respectively. *ΔRatio_Vol* is the first difference in the U.S. trading volume relative to Canada. *ΔRatio_Espread* is the first difference in the U.S. effective spread relative to Canada. *ΔRatio_AT* is the first difference of the U.S. AT activity relative to Canada. Figure in parenthesis is the p-value. *** denotes significance at the 1% level.

| | ΔIS | ΔPT | ΔRatio_Vol | Δ Ratio_Espread | Δ Ratio_AT |
|---|---|---|---|---|---|
| ΔIS | 1 | | | | |
| ΔPT | 0.913*** | 1 | | | |
| | [0.000] | | | | |
| ΔRatio_Vol | 0.161*** | 0.132*** | 1 | | |
| | [0.000] | [0.000] | | | |
| ΔRatio_Espread | −0.109*** | −0.118*** | −0.085*** | 1 | |
| | [0.000] | [0.000] | [0.000] | | |
| ΔRatio_AT | −0.195*** | −0.173*** | −0.670*** | 0.165*** | 1 |
| | [0.000] | [0.000] | [0.000] | [0.000] | |

persistence over time. Once price discovery is gained by a particular market, it tends to stay in that market. The next section analyzes what drives this dynamics in price discovery.

### 5.2. Reduced-form VAR results

We investigate what drives changes in price discovery over time, i.e. how measures of price discovery, liquidity, and AT activity interact with each other. To gain some preliminary insight about the relation between these measures, we test for correlation among them. Table 4 presents the correlation matrix among the VAR variables. Correlation between *ΔIS* and *ΔPT* is 0.913, which shows the high level of similarity between both measures. We observe that *ΔRatio_Vol* is positively correlated with *ΔIS* and *ΔPT*. Both *ΔRatio_Espread* and *ΔRatio_AT* are negatively correlated with *ΔIS* and *ΔPT*. Furthermore, *ΔRatio_AT* is also negatively correlated with *ΔRatio_Vol* and positively correlated with *ΔRatio_Espread*. These coefficients are all highly significant, thus highlighting the importance of the contemporaneous relations among the variables.

To assess the strength and statistical significance of these relations, we start by estimating the reduced-form VAR of Eq. (13) for each of the 38 firms. We report our results in Table 5. Coefficients are the sums of 5-day lagged coefficients, averaged across 38 stocks. The Granger causality test is done per cross-section and the p-value reported (in parentheses) is the average across the 38 stocks.[13]

Panel A and B of Table 5 report the results of the VAR for the IS and PT, respectively. The first column in each panel presents the factors which affect the changes in price discovery measures. We observe that *ΔIS* (*ΔPT*) is positively related to the lagged values of *ΔRatio_Vol* with a coefficient of 0.170 (0.121). A positive change in relative trading volume between the U.S. and Canada over the previous 5 days leads to a positive change in IS (PT) in the following day. This is in line with the argument of Stickel and Verrecchia (1994) that high volume indicates that the demand underlying a price change is informative, and

[13] We also estimate the reduced-form VAR through a multivariate procedure as a pooled VAR. The Granger causality test is also done through a multivariate procedure using all 38 stocks. Overall, the results are comparable with our main findings, and are available on request.

**Table 5**
VAR estimation results. This table presents the sum of the lag coefficients of the VAR in Eq. (13) across the 38 stocks in our sample. The column variable is the dependent variable while the row variable is the explanatory variable. Panel A reports the coefficients from the IS VAR model. Panel B reports the coefficients from the PT VAR model. Figures in brackets are the average p-values from the Granger Causality Test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | Dependent variable | | | |
|---|---|---|---|---|
| | *ΔIS* | *ΔRatio_Vol* | *ΔRatio_Espread* | *ΔRatio_AT* |
| *Panel A: IS reduced-form VAR model* | | | | |
| $\sum \Delta IS_{t-k}$ | −1.907*** | **0.022**** | **− 0.002*** | **− 0.012**** |
| | [0.000] | **[0.027]** | **[0.071]** | **[0.045]** |
| $\sum \Delta Ratio\_Vol_{t-k}$ | **0.170***** | −1.973*** | **− 0.035**** | **− 0.061*** |
| | **[0.009]** | [0.000] | **[0.036]** | **[0.050]** |
| $\sum \Delta Ratio\_Espread_{t-k}$ | **− 0.104**** | **− 0.091**** | −1.868*** | **0.014*** |
| | **[0.033]** | **[0.008]** | [0.000] | **[0.060]** |
| $\sum \Delta Ratio\_AT_{t-k}$ | **− 0.017*** | **− 0.005*** | −0.003 | −1.803*** |
| | **[0.059]** | **[0.085]** | [0.110] | [0.000] |
| Adj. R-squared | 0.30 | 0.29 | 0.32 | 0.27 |
| | Dependent variable | | | |
| | *ΔPT* | *ΔRatio_Vol* | *ΔRatio_Espread* | *ΔRatio_AT* |
| *Panel B: PT reduced-form VAR model* | | | | |
| $\sum \Delta PT_{t-k}$ | −1.873*** | **0.039**** | **− 0.001*** | **− 0.023*** |
| | [0.000] | **[0.035]** | **[0.091]** | **[0.054]** |
| $\sum \Delta Ratio\_Vol_{t-k}$ | **0.121***** | −1.968*** | **− 0.039**** | **− 0.046**** |
| | **[0.008]** | [0.000] | **[0.031]** | **[0.043]** |
| $\sum \Delta Ratio\_Espread_{t-k}$ | **− 0.083*** | **− 0.070**** | −1.855*** | **0.039*** |
| | **[0.075]** | **[0.045]** | [0.000] | **[0.060]** |
| $\sum \Delta Ratio\_AT_{t-k}$ | **− 0.016**** | **− 0.002*** | −0.003 | −1.811*** |
| | **[0.049]** | **[0.085]** | 0.120 | [0.000] |
| Adj. R-squared | 0.30 | 0.29 | 0.32 | 0.27 |

therefore should be incorporated into prices.

We also observe that *ΔIS* (*ΔPT*) is negatively related to lagged *ΔRatio_Espread* with a coefficient of −0.104 (−0.083). A decrease in relative effective spread over the past 5 days leads to a positive change in IS (PT) on the following day. This indicates that as trading costs decrease, price discovery tends to increase, indicating intermarket competition between liquidity providers. This is consistent with the cross-sectional findings of Eun and Sabherwal (2003) who suggest that a lower spread in one market represents a competitive threat faced by liquidity providers in another market. In this case, Canadian liquidity providers become more responsive to U.S. prices.

The impact of *ΔRatio_AT* on *ΔIS* (*ΔPT*) is negative and significant with a coefficient of −0.017 (−0.016). This implies that an increase of AT activity in the U.S. relative to Canada leads to a lower contribution of the U.S. market to price discovery. We interpret this finding as increased AT activity limiting efficiency and inflicting negative congestion externalities on financial markets. High-frequency traders compete aggressively by generating market congestion to slow down one another and create exploitable latency arbitrage opportunities (Gai et al., 2014; Egginton et al., 2016). In the process, they cause a crowding out effect which pushes away investors who are disadvantaged in terms of speed. As a consequence, the price discovery of the market as a whole declines.[14]

The second column in each panel reports the factors which affect the changes in relative trading volume. We observe that lagged values of *ΔIS* (*ΔPT*) have an impact on *ΔRatio_Vol* with a coefficient of 0.022 (0.039), suggesting that improvements in price discovery lead to an increase in relative trading volume. The coefficients of *ΔRatio_Espread* on *ΔRatio_Vol* are negative and significant at −0.091 (−0.070) which suggest that as trading becomes cheaper (relative effective spread decreases), trading volume increases. Furthermore, we find negative coefficients of *ΔRatio_AT* on *ΔRatio_Vol* at −0.074 (−0.079). As relative AT activity increases, relative trading volume decreases. This finding again indicates that greater AT activity has a negative impact and pushes away other traders in the market who are disadvantaged in terms of speed.

The third column shows that there is an impact of lagged values of *ΔIS* (*ΔPT*) on *ΔRatio_Espread* with a magnitude of −0.002 (−0.001). The Granger causality tests show statistically significant results, suggesting that trading costs decrease as a market's contribution to price discovery increases.

The fourth column shows the factors affecting changes in relative AT activity. We observe that the impact of *ΔIS* (*ΔPT*) on *ΔRatio_AT* is negative and significant with a coefficient of −0.012 (−0.023). This finding suggests that a lower contribution to price discovery of one market can provide opportunity to AT. A less informationally efficient market (i.e. lower contribution to price discovery) creates an environment that can potentially be exploited by AT, especially the high-frequency traders who exercise arbitrage strategies to exploit price discrepancies among securities. Consistent with this argument, we find the coefficients of *ΔRatio_Vol* and *ΔRatio_Espread* on *ΔRatio_AT* to be negative and positive, respectively. This further indicates that inefficiencies in the market attract high-frequency trading. Specifically, low volume and wide spreads could suggest less competition among AT and may present greater profits to AT. Hence, algorithmic traders are more willing to trade in a less competitive environment.

The results in Table 5 suggest that relative increases in liquidity (i.e. higher relative trading volume and lower effective spread) lead to a greater contribution of a market to price discovery while an improvement in price discovery leads to greater liquidity. We also observe that an increase in algorithmic trading activity of a market relative to another market leads to lower price discovery. This finding is in line with recent literature on negative externalities of high-frequency trading (Stein, 2009; Gai et al., 2014; Egginton et al., 2016). In particular, we find that although high-frequency trading has improved the informativeness of quotes and price discovery of the faster traders (Hendershott et al., 2011; Riordan & Storkenmaier, 2012; Brogaard et al., 2017), their increased speed and information processing abilities discourage other traders from submitting limit orders. As a result, aggressive investment strategies by the high-frequency traders lead to a crowding out effect that pushes slower traders away from the market, resulting in lower price discovery.[15]

### 5.3. Structural VAR results

In addition to lagged effects, we also assess the contemporaneous causal relations between variables using the identification through heteroskedasticity approach (Rigobon, 2003). The structural parameters in Eq. (12) are estimated by GMM for each of the 38 firms separately. The coefficients are then averaged while the standard errors

[14] In addition, Abergel, Bouchaud, Foucault, Lehalle, and Rosenbaum (2012) explain that high-frequency traders, in some cases, use their speed advantage to free-ride on trade-related information (e.g. order flow, prices, volume, duration between trades) acquired by informed investors. This may reduce investors' incentives to acquire information in the first place, leading to lower price discovery.

[15] It is important to note that the Canadian market is fragmented considerably after 2009 with the opening of alternative trading venues such as Alpha Trading, Chi-X and Pure Trading. Such fragmentation could affect our results. We test the robustness of our findings by reconstructing the variables $Vol^{CAN\_ALL}$, $Espread^{CAN\_ALL}$, and $AT^{CAN\_ALL}$. $Vol^{CAN\_ALL}$ and $Espread^{CAN\_ALL}$ are calculated as the total volume traded and the volume-weighted effective spread across all the exchanges, respectively. $AT^{CAN\_ALL}$ is calculated based on the total messages submitted and the total dollar trading volume across all the Canadian exchanges. The reduced-form VAR of Eq. (13) and the structural VAR of Eq. (12) are then re-estimated using the aggregated data. We observe that the coefficients remain identical to those reported in Tables 5 and 6, suggesting that the dynamic relationships among the variables in the VAR are not driven by the fragmentation in the Canadian financial market. A discussion on this fragmentation and results can be found in Appendix B.

**Table 6**
Contemporaneous relation between variables. This table presents the coefficients for the contemporaneous interactions between the VAR variables across the 38 stocks in our sample. Note that the coefficients in this table have the opposite signs to the coefficients of matrix $A$ because matrix $A$ is on the left-hand side of Eq. (12). When taken to the right-hand side the effects become positive. Subsequently, the column variable is the dependent variable while the row variable is the explanatory variable. Panel A reports the results from the IS VAR model. Panel B reports the results from the PT VAR model. Figures in brackets are the average p-values. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\Delta IS$ | $\Delta Ratio\_Vol$ | $\Delta Ratio\_Espread$ | $\Delta Ratio\_AT$ |
| *Panel A: IS structural VAR model* | | | | |
| $\Delta IS_t$ | 1 | **0.024*****<br>**[0.003]** | **− 0.007***<br>**[0.071]** | **− 0.044*****<br>**[0.000]** |
| $\Delta Ratio\_Vol_t$ | 0.032<br>[0.113] | 1 | 0.005<br>[0.417] | **− 0.323*****<br>**[0.000]** |
| $\Delta Ratio\_Espread_t$ | **− 0.246*****<br>**[0.000]** | **− 0.062***<br>**[0.053]** | 1 | **0.123*****<br>**[0.000]** |
| $\Delta Ratio\_AT_t$ | **− 0.122*****<br>**[0.000]** | **− 0.456*****<br>**[0.000]** | **0.037*****<br>**[0.003]** | 1 |
| | Dependent variable | | | |
| | $\Delta PT$ | $\Delta Ratio\_Vol$ | $\Delta Ratio\_Espread$ | $\Delta Ratio\_AT$ |
| *Panel B: PT structural VAR model* | | | | |
| $\Delta PT_t$ | 1 | 0.001<br>[0.934] | **− 0.017*****<br>**[0.003]** | **− 0.027*****<br>**[0.003]** |
| $\Delta Ratio\_Vol_t$ | **0.030***<br>**[0.086]** | 1 | 0.006<br>[0.235] | **− 0.301*****<br>**[0.000]** |
| $\Delta Ratio\_Espread_t$ | **− 0.203****<br>**[0.000]** | **− 0.042***<br>**[0.091]** | 1 | **0.114*****<br>**[0.001]** |
| $\Delta Ratio\_AT_t$ | **− 0.125*****<br>**[0.000]** | **− 0.494*****<br>**[0.000]** | **0.049*****<br>**[0.000]** | 1 |

are computed cross-sectionally.

Table 6 reports the results for the contemporaneous relation between the variables in the structural VAR model with results for IS in Panel A and PT in Panel B. The first column of each panel reports the impact of liquidity and AT activity on price discovery. We observe a significant and positive causal effect of $\Delta Ratio\_Vol$ on $\Delta PT$ with a coefficient of 0.030. There is a strong negative contemporaneous effect of $\Delta Ratio\_Espread$ on $\Delta IS$ ($\Delta PT$) with a coefficient of −0.246 (−0.203). The last row of each Panel indicates a negative contemporaneous interaction of $\Delta Ratio\_AT$ on $\Delta IS$ ($\Delta PT$) at −0.122 (−0.125). The fact that these relations are observed in both structural and reduced-form VAR models suggests that liquidity and AT activity affect price discovery instantaneously as well as with a lag.

The second column reports the coefficients for the determinants of $\Delta Ratio\_Vol$. We observe a significant positive relation between $\Delta IS$ and $\Delta Ratio\_Vol$, indicating that price discovery is related to trading volume. We also find negative relation between $\Delta Ratio\_Espread$ and $\Delta Ratio\_Vol$. Furthermore, the contemporaneous impact of $\Delta Ratio\_AT$ on $\Delta Ratio\_Vol$ is negative and highly significant at −0.456 (−0.494), indicating that the impact of AT on relative trading volume is more prevalent contemporaneously, i.e. as high-frequency traders enter the market, trading activity by the slower traders decreases.

In the third column, we observe that $\Delta IS$ ($\Delta PT$) negatively affects $\Delta Ratio\_Espread$ with a coefficient of −0.007 (−0.017), suggesting that an increase in price discovery leads to a decrease in relative spread. We also observe that $\Delta Ratio\_AT$ significantly affects $\Delta Ratio\_Espread$, which was not observed in Table 5. We interpret this as AT pushing away other traders in the market who are relatively disadvantaged in terms of speed, hence causing the spread to increase.

Finally, in the last column, we observe similar significant relations as previously observed in Table 5. However, the coefficients of $\Delta Ratio\_Vol$ on $\Delta Ratio\_AT$ and of $\Delta Ratio\_Espread$ on $\Delta Ratio\_AT$ are greater in magnitude at −0.323 (−0.301) and 0.123 (0.114) for the IS (PT) model, respectively. These results suggest that AT activity reacts strongly to changes in liquidity within the same day. Overall, our results in Table 6 show that there exists not only lagged, but also contemporaneous relations among relative liquidity, AT activity, and price discovery.[16]

### 5.4. Price discovery dynamics and the order protection rule

In a further test, we assess the impact of regulatory changes in the U.S. and Canadian markets. The Order Protection Rule (OPR) requires that marketplaces enforce policies to ensure consistent price quotation and prevent trading through a better priced order on another market. In the U.S. it was implemented on 8 October 2007 as part of the Reg. NMS which was prompted by the Securities and Exchange Commission to improve fairness in price execution. In Canada, the OPR was implemented by the Canadian Securities Administrators (CSA) on 1 February 2011.

With regards to Reg. NMS, there are several studies assessing market quality following the new regulation. For instance, Hendershott and Jones (2005) suggest that an increase in market fragmentation leads to slower price discovery. Hence, regulatory changes to create a more integrated market should improve price discovery. Furthermore, Barclay, Hendershott, and Jones (2008) find that the consolidation of orders is important for producing efficient prices, especially during times of high liquidity demand. On the contrary, Chung and Chuwonganant (2012) examine the liquidity of the U.S. stock markets 1 month before and after the adoption of Reg. NMS and find that liquidity was reduced in the form of increased quoted and effective spreads, as well as decreased quoted dollar depth. This evidence indicates that there may be an impact of Reg. NMS (and the OPR) on the dynamics of price discovery.

In this section, we first show how price discovery, liquidity, and AT activity changed after the implementation of the OPR in both markets. We then examine whether the new regulation affects the dynamics of price discovery for cross-listed stocks. To this end, we split our data into three sub-periods. The first sub-period is from 1 January 2004 to 7 October 2007 as the pre-OPR period in the U.S. The second sub-period is from 8 October 2007 to 31 January 2011 as the post-OPR period in the U.S. (and the pre-OPR period in Canada). The third sub-period is from 1 February 2011 to 31 August 2017 as the post-OPR period in Canada.[17]

Table 7 reports the percentage change in price discovery, liquidity, and AT measures between pre- and post-OPR periods. Panel A reports the change over the first sub-period to the second sub-period. We observe that trading volume in the U.S. increased significantly by 279% compared to Canada where it only increased by 84%. This finding indicates a much larger increase in liquidity in the U.S. compared to Canada after the new regulation. Consequently, relative trading volume increased by 78%. Effective spreads, on the other hand, did not change significantly in either markets. Contrary to Chung and Chuwonganant (2012), we do not observe an increase in spreads after the adoption of Reg. NMS, but rather an improvement in trading volume. As for AT activity, the U.S. market experienced a significant increase of 40%. In Canada, the increase in AT activity is more substantial at 69%. These findings are in line with Panel C of Fig. 1, which shows that the increase in AT activity is much higher in Canada than in the U.S. We observe that both IS and PT increased significantly by 97% and 53%, respectively, suggesting that the U.S. contribution to price discovery has increased significantly after the new regulation in the U.S. These findings are in line with Hendershott and Jones (2005) and Barclay et al. (2008)

[16] Similar to previous section, we provide a robustness test using aggregated data across the Canadian exchanges. The results can be found in Appendix B.

[17] We also conducted our analysis using the period 1 January 2004 to 20 May 2007 (the start of the pilot stocks phase) as the pre-Reg. NMS period in the U.S. We do not find substantially different results than those reported in Tables 7 and 8.

**Table 7**
Change in variables surrounding the implementation of the order protection rule. This table provides the change in price discovery, liquidity, and algorithmic trading activity measures for 38 Canadian cross-listed stocks. The figures reported are the percentage differences before and after the adoption of OPR in the U.S. on 8 October 2007 and in Canada on 1 February 2011. Figures in parentheses are the t-statistics. **, and *** denote significance at the 5%, and 1% levels, respectively.

| | Panel A: | | Panel B: | |
|---|---|---|---|---|
| | Pre and Post US OPR | | Pre and Post CAN OPR | |
| | Diff | t-stat | Diff | t-stat |
| $Vol^{US}$ | 279%*** | (7.52) | −5% | (−0.63) |
| $Vol^{CAN}$ | 84%*** | (4.51) | 33%** | (2.23) |
| *Ratio_Vol* | 78%*** | (5.00) | −7% | (−1.40) |
| $Espread^{US}$ | −4% | (−0.31) | 11% | (0.83) |
| $Espread^{CAN}$ | −10% | (−0.89) | −4% | (−0.35) |
| *Ratio_Espread* | 3%** | (2.02) | 8%*** | (8.27) |
| $AT^{US}$ | 40%*** | (7.21) | 20%*** | (−2.58) |
| $AT^{CAN}$ | 69%*** | (33.53) | 53%*** | (−7.95) |
| *Ratio_AT* | −19%*** | (−8.68) | −12%*** | (−3.31) |
| *IS* | 97%*** | (8.50) | −4% | (−1.53) |
| *PT* | 53%*** | (10.75) | −3% | (−1.04) |

**Table 8**
Sub-periods VAR estimation results. This table presents the IS reduced-form VAR results at three sub-periods surrounding the implementation of the Order Protection Rule in the U.S. and Canada. Panel A reports the results for the period from 1 January 2004 to 7 October 2007, Panel B from 8 October 2007 to 31 January 2011 and Panel C from 1 February 2011 to 31 August 2017. The column variable is the dependent variable while the row variable is the explanatory variable. Figures in brackets are the average p-values from the Granger Causality Test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\Delta IS$ | $\Delta Ratio\_Vol$ | $\Delta Ratio\_Espread$ | $\Delta Ratio\_AT$ |
| *Panel A: IS reduced-form VAR model (Jan 2004–Oct 2007)* | | | | |
| $\sum \Delta IS_{t-k}$ | −2.202*** | 0.012** | − 0.003** | −0.024 |
| | [0.000] | [0.018] | [0.041] | [0.409] |
| $\sum \Delta Ratio\_Vol_{t-k}$ | 0.160*** | −1.881*** | −0.024 | − 0.107*** |
| | [0.003] | [0.000] | [0.342] | [0.000] |
| $\sum \Delta Ratio\_Espread_{t-k}$ | − 0.095* | − 0.091** | −2.096*** | 0.002** |
| | [0.071] | [0.018] | [0.000] | [0.021] |
| $\sum \Delta Ratio\_AT_{t-k}$ | − 0.088** | − 0.100*** | 0.007 | −1.902*** |
| | [0.035] | [0.000] | [0.819] | [0.000] |
| Adj. R-squared | 0.37 | 0.31 | 0.34 | 0.29 |
| *Panel B: IS reduced-form VAR model (Oct 2007–Jan 2011)* | | | | |
| $\sum \Delta IS_{t-k}$ | −2.025*** | 0.069*** | 0.0035 | − 0.039*** |
| | [0.000] | [0.000] | [0.276] | [0.003] |
| $\sum \Delta Ratio\_Vol_{t-k}$ | 0.157*** | −1.886*** | 0.006 | − 0.076** |
| | [0.000] | [0.000] | [0.982] | [0.037] |
| $\sum \Delta Ratio\_Espread_{t-k}$ | − 0.237*** | − 0.329*** | −1.913*** | 0.143 |
| | [0.001] | [0.000] | [0.000] | [0.145] |
| $\sum \Delta Ratio\_AT_{t-k}$ | − 0.026*** | − 0.065*** | − 0.010* | −1.750*** |
| | [0.005] | [0.000] | [0.092] | [0.000] |
| Adj. R-squared | 0.34 | 0.30 | 0.32 | 0.27 |
| *Panel C: IS reduced-form VAR model (Feb 2011–Aug 2017)* | | | | |
| $\sum \Delta IS_{t-k}$ | −1.974 | 0.063*** | − 0.004* | − 0.022** |
| | | [0.000] | [0.060] | [0.018] |
| $\sum \Delta Ratio\_Vol_{t-k}$ | 0.016** | −1.917 | − 0.007*** | −0.022 |
| | [0.043] | | [0.002] | [0.435] |
| $\sum \Delta Ratio\_Espread_{t-k}$ | −0.051 | − 0.143** | −2.076 | 0.053 |
| | [0.158] | [0.016] | | [0.118] |
| $\sum \Delta Ratio\_AT_{t-k}$ | − 0.004** | − 0.131*** | 0.013** | −1.607 |
| | [0.014] | [0.000] | [0.035] | |
| Adj. R-squared | 0.34 | 0.31 | 0.35 | 0.27 |

who advocate that a new regulation to create a more integrated market would lead to greater price discovery. Panel B reports the change over the second sub-period to the third sub-period, during the implementation of the OPR in Canada. We do not observe changes in U.S. trading activity, but in Canada, trading volume has increased by 33%. We do not find that effective spreads change significantly in either markets. AT activity in the U.S. increased by 20%, whereas in Canada, it increased more substantially by 53%. Both the IS and PT do not change significantly following the OPR in Canada, suggesting that the U.S. contribution to price discovery was not affected by the implementation of

**Table 9**
Sub-periods contemporaneous relation results. This table presents the IS structural VAR results at three sub-periods surrounding the implementation of the Order Protection Rule in the U.S. and Canada. Panel A reports the results for the period from 1 January 2004 to 7 October 2007, Panel B from 8 October 2007 to 31 January 2011 and Panel C from 1 February 2011 to 31 August 2017. Note that the coefficients in this table have the opposite signs to the coefficients of matrix $A$ because matrix $A$ is on the left-hand side of Eq. (12). When taken to the right-hand side the effects become positive. Subsequently, the column variable is the dependent variable while the row variable is the explanatory variable. Figures in brackets are the average p-values. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\Delta IS$ | $\Delta Ratio\_Vol$ | $\Delta Ratio\_Espread$ | $\Delta Ratio\_AT$ |
| *Panel A: IS structural VAR model (Jan 2004–Oct 2007)* | | | | |
| $\Delta IS_t$ | 1 | −0.002 | −0.006 | **− 0.032**** |
| | | [0.832] | [0.335] | **[0.016]** |
| $\Delta Ratio\_Vol_t$ | **0.123*** | 1 | **0.067**** | **− 0.392***** |
| | **[0.084]** | | **[ 0.010]** | **[0.000]** |
| $\Delta Ratio\_Espread_t$ | **− 0.327***** | 0.012 | 1 | **0.278***** |
| | **[0.002]** | [0.772] | | **[0.000]** |
| $\Delta Ratio\_AT_t$ | **− 0.175**** | **− 0.590***** | **0.155***** | 1 |
| | **[0.016]** | **[0.000]** | **[0.000]** | |
| *Panel B: IS structural VAR model (Oct 2007–Jan 2011)* | | | | |
| $\Delta IS_t$ | 1 | 0.023 | **− 0.014*** | **− 0.032*** |
| | | [0.309] | **[0.078]** | **[0.060]** |
| $\Delta Ratio\_Vol_t$ | **0.087**** | 1 | −0.015 | **− 0.359***** |
| | **[0.020]** | | [0.100] | **[0.000]** |
| $\Delta Ratio\_Espread_t$ | **− 0.150**** | −0.054 | 1 | **0.230***** |
| | **[0.037]** | [0.349] | | **[0.001]** |
| $\Delta Ratio\_AT_t$ | **− 0.057*** | **− 0.430***** | −0.006 | 1 |
| | **[0.090]** | **[0.000]** | [0.657] | |
| *Panel C: IS structural VAR model (Feb 2011–Aug 2017)* | | | | |
| $\Delta IS_t$ | 1 | 0.022 | **− 0.024***** | −0.022 |
| | | [0.115] | **[0.000]** | [0.137] |
| $\Delta Ratio\_Vol_t$ | 0.061 | 1 | −0.006 | **− 0.281***** |
| | [0.136] | | [0.201] | **[0.000]** |
| $\Delta Ratio\_Espread_t$ | −0.083 | −0.086 | 1 | 0.021 |
| | [0.112] | [0.350] | | [0.670] |
| $\Delta Ratio\_AT_t$ | **− 0.177***** | **− 0.507***** | **0.013*** | 1 |
| | **[0.000]** | **[0.000]** | **[0.076]** | |

the OPR in Canada.

We test the impact of the OPRs on price discovery dynamics by examining the relations between liquidity, AT activity and price discovery measures during the three sub-periods. Table 8 shows the result of the VAR analysis of Eq. (13) for the three sub-periods: before the adoption of Reg. NMS, after Reg. NMS but before the introduction of the OPR in Canada, and after the introduction of the OPRs. For brevity, we only report the results from the IS VAR model.[18] Overall, we do not observe any significant differences from the results reported in Table 5. As shown in Panel A, B and C of Table 8, changes in relative trading volume positively affect the changes in IS as shown by the highly significant p-values from the Granger causality tests. We also observe that changes in relative effective spread and relative AT activity are negatively related to changes in IS. In the opposite direction, we observe that changes in IS lead to positive changes in relative trading volume as shown by the first row of the third column in each Panel. The impact on changes in relative effective spread remains small and significant for the first sub-period, but insignificant for the second sub-period. The coefficients for the changes in relative AT activity are also negative, despite being significant only in the second sub-period. Based on these observations, we conclude that while the OPR had a significant impact on the trading environment in the U.S. and affected the level of price discovery, it did not affect the underlying relations between price discovery, liquidity and algorithmic trading activity.

Table 9 shows the contemporaneous relations of the VAR variables in Eq. (12) during the three sub-periods. Similar to the results in Table 6, we observe a uni-directional relation between liquidity measures and price discovery. Specifically, changes in relative trading volume contemporaneously and positively affect the changes in IS, while changes in relative effective spread contemporaneously and negatively affect the changes in IS. The bi-directional negative relation between AT activity and price discovery measures still persists. Overall, both Tables 8 and 9 show that the relations between price discovery and liquidity and AT measures persist even after taking into account the regulatory changes in both the U.S. and Canadian financial markets.

## 6. Conclusion

In this paper, we study price discovery dynamics for a sample of Canadian cross-listed stocks in the U.S. from January 2004 to August 2017. We compute daily measures of price discovery and assess the causal relations between price discovery, liquidity, and algorithmic trading activity. To accommodate for both lagged and contemporaneous relations among the variables, we follow the approach of Chaboud et al. (2014) by estimating a reduced-form VAR, as well as a structural VAR using the identification through heteroskedasticity approach developed by Rigobon (2003).

We show that price discovery of the U.S. market relative to Canada exhibits an upward trend, suggesting that over time, the U.S. market is becoming more dominant in terms of price discovery for Canadian cross-listed stocks. Assessing the dynamics involved, we find that liquidity is related to price discovery. Improvements in relative liquidity (an increase in trading volume and a decrease in effective spread in one market relative to another) increase that market's contribution to price discovery. This finding implies that the market which provides better liquidity will become more important in terms of price discovery. This impact occurs instantaneously (within the same day) as well as with a protracted lag (after several days). Conversely, we find that an increase in price discovery leads to improved liquidity, indicating that the market which leads in terms of price discovery attracts more liquidity. We also find that relative algorithmic trading activity is negatively related to price discovery, which is consistent with the recent literature on negative externalities of high-frequency trading. Particularly, as high-frequency traders compete aggressively with one another to create latency arbitrage opportunity, they cause a crowding-out effect which pushes away investors who are disadvantaged in terms of speed, degrading market quality. We further observe that while the U.S. market's contribution to price discovery increased after the adoption of the OPR, the dynamic relations between price discovery, liquidity and algorithmic trading activity do not change, lending support to the robustness of our results.

## Appendix A. Identification through heteroskedasticity

This appendix discusses the identification through heteroskedasticity methodology as employed in our study. Rigobon (2003) provides the theoretical derivation of the methodology.

The identification through heteroskedasticity methodology is used when measuring contemporaneous relationships among the variables in a structural VAR model. Recall the SVAR equation,

[18] The PT VAR model yields similar results and these are available upon request.

$$A\Delta Y_t = c + \sum_{k=1}^{K} \Pi_k \Delta Y_{t-k} + \varepsilon_t, \tag{A.1}$$

and the reduced-form VAR,

$$\Delta Y_t = A^{-1}c + A^{-1} \sum_{k=1}^{K} \Pi_k \Delta Y_{t-k} + A^{-1}\varepsilon_t. \tag{A.2}$$

To assess the contemporaneous interactions of our model, we are interested in estimating the parameters in matrix $A$, which represents the structural parameters linking the variables to each other. There is one issue, however, that is the structural form of Eq. (A.1) cannot be estimated directly using OLS, although the reduced-form of Eq. (A.2) can. Given the relation between the structural and the reduced-form, we can therefore obtain the parameters in matrix $A$ by first estimating the reduced-form model.

In our empirical setting, we obtain the parameters in $A$ through the following procedure. First, we estimate the reduced-form VAR in Eq. (A.2) using OLS. The lag specification is determined by the Schwartz Information Criterion, which in our case suggests an optimal lag-length of 5 days. From this step, we obtain the reduced-form residuals $\tilde{\varepsilon}_t$, which contain the contemporaneous effects.

Second, from the reduced-form residuals, we define the heteroskedastic regimes. We do so by computing rolling window variances of 20 observations each, following Ehrmann et al. (2011). A regime is identified if one variance of a variable exceeds the average variance of that variable over the sample period plus one standard deviation, while at the same time the variances of the other three variables do not exceed their average variances plus one standard deviation.[19] Using this approach, we identify 6 regimes in total: 1 regime to represent a tranquil state where all the four variables do not exhibit elevated conditional volatility; 4 regimes where only one variable exhibits elevated conditional volatility while the other three are stable; and 1 regime where at least 2 variables exhibit elevated conditional volatility.

Third, once the regimes are identified, we estimate the variance-covariance matrices, $\widetilde{\Omega}_s$, of the reduced-form residuals in variance regime $s$ ($s = 1,2,\ldots,6$). Given that $\Omega_s$ are the variance-covariance matrices of the SVAR that we are interested in, and assuming the following moment conditions hold,

$$A\widetilde{\Omega}_s A' = \Omega_s. \tag{A.3}$$

The parameters in $A$ and $\Omega_s$ can be estimated using GMM by minimizing the following objective function:

$$\min g'g \quad \text{with} \quad g = A\widetilde{\Omega}_s A' - \Omega_s. \tag{A.4}$$

Identification is achieved as long as the covariance matrices constitute a system of equations that is linearly independent. Therefore, the basic idea of the identification through heteroskedasticity approach is to increase the number of available moments or equations and obtain matrix A which satisfies Eq. (A.3) across different regime, $s$. If the variance of the shocks in the system changes across different regimes, but the parameters in matrix $A$ remain constant, the system may be identified. Hence, the shift in variances provides an extra source of variation needed for identification in the presence of endogeneity, such that matrix $A$ can be estimated.

There are several important points to note. First, the structural shocks across different regimes need to be uncorrelated. If the shocks were correlated and there is no restriction on the variation of such covariance, then every heteroskedastic regime adds as many equations as unknowns, and the problem of identification cannot be solved. The zero correlation assumption, therefore, needs to be maintained for the identification to work.

Second, the parameters in matrix $A$ needs to be stable across the different volatility regimes. This means that even though exogenous shocks may have changed the conditional volatility, the structural relationships need to remain constant. This is slightly harder to interpret because while in some instances one can point to an event that is likely to have affected the variances of the shocks, in most cases, the shift in variances is not necessarily associated with a specific event. Thus, heteroskedasticity in the data can still be observed in the absence of a structural event. As such, a more pragmatic interpretation is to view the results as conditional on the structural relationship remaining constant. That is, provided the relationship is constant over time, the method delivers consistent estimates of the contemporaneous effects.

Table A.1 shows the variances of the reduced-form VAR residuals for $Y_t$, across the 6 different heteroskedastic regimes as well as the frequency of occurences of those regimes. As shown, Regime 1 (tranquil state) has the highest number of occurrence at 68% and Regime 6 (elevated state) at 10%. The remaining regimes occur at around 5%–6% of the time. Comparing the benchmark variances in regime 1 with those in regimes 2 to 6, it is clear that the variances are no longer proportional to each other. Specifically, in the case of regime 2 to 5, one variable exhibits elevated conditional volatility and the other three are stable, while in the case of regime 6, all 4 variables exhibit elevated conditional volatility. The existence of heteroskedastic regimes solves the identification problem as they increase the number of Eq. (A.3). This is an important identification requirement, especially to disentangle the dynamics of each of the variables in the VAR, which normally are highly correlated. As with other studies using this identification approach, the precise form of the heteroskedasticiy is of no particular interest (see e.g. Ehrmann et al., 2011; Badshah et al., 2013; Chaboud et al., 2014). What matters is that the estimates of the coefficients are consistent, regardless of how the heteroskedasticity is modeled.

[19] To assess the sensitivity of the results with our regime definition, we consider two other thresholds. First, we define a regime if the variance of a variable exceeds the average variance of that variable over the sample period plus one and a half standard deviation (compared with the one standard deviation in variance), while at the same time, the variances of the other three variables do not exceed their average variances plus one and a half standard deviation. Second, we define a regime if the variance of a variable exceeds the average variance of that variable over the sample period plus 0.75 standard deviation, while at the same time, the variances of the other three variables do not exceed their average variances plus 0.75 deviation. The results based on these regime definitions are very similar to our findings in Table A.1.

Table A.1
Variances of the reduced-form VAR residuals. This table reports the variances of the residuals from the reduced-form VAR in Eq. (A.2). The residuals are split into 6 different regimes based on various elevations in conditional volatility. $R1$ represents a tranquil state where all four variables do not exhibit elevated conditional volatility. $R2$ to $R5$ represent the regimes where only one variable exhibits elevated conditional volatility while the other three are stable. $R6$ represents a regime where at least 2 variables exhibit elevated conditional volatility. The first figure in each column represents the average variance across our sample. The second figure (in parantheses) represents the percentage change from the variance in the tranquil state, $R1$. The last row in each panel report the frequency of occurrences of the various regimes.

| | $R1$ | $R2$ | | $R3$ | | $R4$ | | $R5$ | | $R6$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: IS structural VAR model* | | | | | | | | | | | |
| Δ*IS* | 0.64 | 3.00 | (366%) | 0.71 | (11%) | 1.01 | (56%) | 0.92 | (42%) | 2.39 | (271%) |
| Δ*Ratio_Vol* | 0.52 | 0.65 | (24%) | 1.60 | (206%) | 0.52 | (0%) | 0.74 | (42%) | 1.26 | (141%) |
| Δ*Ratio_Espread* | 0.07 | 0.14 | (82%) | 0.08 | (9%) | 0.42 | (468%) | 0.10 | (29%) | 0.25 | (235%) |
| Δ*Ratio_AT* | 0.40 | 0.49 | (23%) | 0.51 | (29%) | 0.43 | (8%) | 1.15 | (188%) | 1.02 | (155%) |
| %obs | 68% | 5% | | 6% | | 5% | | 6% | | 10 % | |
| *Panel B: PT structural VAR model* | | | | | | | | | | | |
| Δ*PT* | 0.58 | 2.69 | (362%) | 0.61 | (5%) | 0.86 | (48%) | 0.77 | (32%) | 2.02 | (248%) |
| Δ*Ratio_Vol* | 0.53 | 0.60 | (13%) | 1.61 | (203%) | 0.54 | (2%) | 0.79 | (50%) | 1.24 | (134%) |
| Δ*Ratio_Espread* | 0.07 | 0.13 | (78%) | 0.08 | (14%) | 0.46 | (531%) | 0.11 | (52%) | 0.22 | (206%) |
| Δ*Ratio_AT* | 0.40 | 0.46 | (15%) | 0.52 | (31%) | 0.41 | (1%) | 1.11 | (176%) | 0.98 | (145%) |
| %obs | 65% | 5% | | 6% | | 5% | | 6% | | 10% | |

## Appendix B. Fragmentation in the Canadian financial market

The Canadian market fragments considerably after 2009 with the opening of alternative trading venues such as Alpha Trading (ALP), Chi-X (CXC) and Pure Trading (GO) among other smaller ones.[20] These exchanges were fragmented due to the absence of the consolidated tape and the Order Protection Rule. As a consequence, dealers were free to decide where to trade and could carry out orders at a non-optimal price even though a better price was available on the same exchange or other exchanges. Such fragmentation may explain why price discovery is more likely to take place in the U.S and could affect the validity of our results, particularly during the period prior to February 2011 when the Order Protection Rule was not yet put in place and consolidated data were unavailable in Canada.

In this section, we provide robustness tests to ensure that our findings are not affected by the fragmentation of the Canadian market. We focus on the four largest Canadian exchanges (TSX, ALP, CXC, GO) which, at the time, account for 98.1% of all trades that occur in Canada. For the period January 2004 to January 2011, we collect trade and quote data from these exchanges, and use them to reconstruct our VAR variables, $Vol^{CAN\_ALL}$, $Espread^{CAN\_ALL}$, and $AT^{CAN\_ALL}$. $Vol^{CAN\_ALL}$ and $Espread^{CAN\_ALL}$ are calculated as the total volume traded and the volume-weighted effective spread across all the exchanges, respectively. $AT^{CAN\_ALL}$ is calculated based on the total messages submitted and total dollar trading volume across all the exchanges. The reduced-form VAR of Eq. (13) and the structural VAR of Eq. (12) are then re-estimated using the aggregated data. If fragmentation does affect the dynamics relationship among our variables, then we can expect our new VAR coefficients to differ from those reported in Tables 5 and 6.

Table B.1 reports the coefficients from the reduced-form and structural VAR models. The re-estimated coefficients remain identical to our findings thus far, suggesting that the dynamic relationships among the variables in the VAR hold and are not affected by the fragmentation in the Canadian financial market.

[20] Clark (2011) finds that in the year 2010, the TSX accounts for 66% of the total trading volume on all Canadian-listed issues, while Alpha Trading was 23.4%, Chi-X was 6.5%, Pure Trading was 2.2%, and the remaining 1.9% was shared among MATCH Now, Omega, and Liquidnet.

Table B.1
VAR Coefficients accounting for market fragmentation in Canada. This table presents the coefficients of the VAR variables after taking into account the fragmentation in the Canadian financial market. The sample period is from January 2004 to January 2011. Panel A reports the results from the IS reduced-form VAR model. Panel B reports the results from the IS structural VAR model. Figures in brackets are the p-values. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\Delta IS$ | $\Delta Ratio\_Vol$ | $\Delta Ratio\_Espread$ | $\Delta Ratio\_AT$ |
| *Panel A: IS reduced-form VAR model* | | | | |
| $\Delta IS_{t-k}$ | −2.141 | **0.044*** | **− 0.012*** | **− 0.047*** |
| | [0.000] | **[0.000]** | **[0.067]** | **[0.000]** |
| $\Delta Ratio\_Vol_{t-k}$ | **0.089*** | −1.763 | **− 0.052*** | **− 0.300*** |
| | **[0.002]** | [0.000] | **[0.012]** | **[0.000]** |
| $\Delta Ratio\_Espread_{t-k}$ | **− 0.088*** | **− 0.036*** | −1.955 | **0.023*** |
| | **[0.047]** | **[0.005]** | [0.000] | **[0.039]** |
| $\Delta Ratio\_AT_{t-k}$ | **− 0.094*** | **− 0.073*** | **− 0.026*** | −1.491 |
| | **[0.059]** | **[0.000]** | **[0.082]** | [0.000] |
| Adj. R-squared | 0.35 | 0.31 | 0.33 | 0.29 |
| *Panel B: IS structural VAR model* | | | | |
| $\Delta IS_t$ | 1 | **0.034*** | −0.004 | **− 0.029*** |
| | | **[0.024]** | [0.309] | **[0.024]** |
| $\Delta Ratio\_Vol_t$ | **0.092*** | 1 | −0.020 | **− 0.489*** |
| | **[0.020]** | | [0.262] | **[0.000]** |
| $\Delta Ratio\_Espread_t$ | **− 0.106*** | **− 0.063*** | 1 | **0.080*** |
| | **[0.000]** | **[0.011]** | | **[0.000]** |
| $\Delta Ratio\_AT_t$ | **− 0.123*** | **− 0.369*** | **0.088*** | 1 |
| | **[0.003]** | **[0.000]** | **[0.000]** | |

## References


Abergel, F., Bouchaud, J. P., Foucault, T., Lehalle, C. A., & Rosenbaum, M. (2012). *Market microstructure: Confronting many viewpoints west sussex*. United Kingdom: John Wiley and Sons.
Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies, 1*(1), 3–40.
Badshah, I. U., Frijns, B., & Tourani-Rad, A. (2013). Contemporaneous spill-over among equity, gold, and exchange rate implied volatility indices. *Journal of Futures Markets, 33*(6), 555–572.
Barclay, M. J., Hendershott, T., & Jones, C. M. (2008). Order consolidation, price efficiency, and extreme liquidity shocks. *Journal of Financial and Quantitative Analysis, 43*(1), 93–122.
Beveridge, S., & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics, 7*, 151–174.
Biais, B., Foucault, T., & Moinas, S. (2015). Equilibrium fast trading. *Journal of Financial Economics, 116*, 292–313.
Boehmer, E., Fong, K., & Wu, J. (2015). *International evidence on algorithmic trading. Working paper*.
Brogaard, J. A. (2010). *High frequency trading and its impact on market quality. Working paper*. Northwestern University.
Brogaard, J., Hendershott, T., & Riordan, R. (2017). *Price discovery without trading: Evidence from limit orders. Working paper*.
Clark, R. (2011). The evolving regulatory landscape in Canada. *Investment Technology Group Insights Publication, 7*(7).
Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *Journal of Finance, 69*(5), 2045–2084.
Chen, H., & Choi, P. M. S. (2012). Does information vault Niagara Falls? cross-listed trading in new york and toronto. *Journal of Empirical Finance, 19*(2), 175–199.
Chung, K. H., & Chuwonganant, C. (2012). Regulation NMS and market quality. *Financial Management, 41*(2), 285–317.
Egginton, J. F., Ness, B. F. V., & Ness, R. A. V. (2016). Quote stuffing. *Financial Management, 45*(3), 583–608.
Ehrmann, M., Fratzscher, M., & Rigobon, R. (2011). Stocks, bonds, money markets and exchange rates: Measuring international financial transmission. *Journal of Applied Econometrics, 26*, 948–974.
Eun, C. S., & Sabherwal, S. (2003). Cross-border listings and price discovery: Evidence from U.S.-listed Canadian stocks. *Journal of Finance, 58*(2), 549–575.
Fleming, J., Ostdiek, B., & Whaley, R. E. (1996). Trading costs and the relative rates of price discovery in stock, futures, and option markets. *Journal of Futures Markets, 16*, 353–387.
Frijns, B., Gilbert, A., & Tourani-Rad, A. (2010). The dynamics of price discovery for cross-listed shares: Evidence from Australia and New Zealand. *Journal of Banking and Finance, 34*, 498–508.
Frijns, B., Gilbert, A., & Tourani-Rad, A. (2015). The determinants of price discovery: Evidence from US-Canadian cross-listed shares. *Journal of Banking and Finance, 59*, 457–468.
Gai, J., Yao, C., & Ye, M. (2014). *The externalities of High-Frequency trading. Working paper*. University of Illinois at Urbana-Champaign.
Gibbs, C. M. (2007). *Adoption of electronic systems and algorithmic trading tools is on the rise with Canadian traders*. WallStreet & Technology Retrieved from http://www.wallstreetandtech.com/trading-technology/adoption-of-electronic-systems-and-algorithmic-trading-tools-is-on-the-rise-with-canadian-traders/d/d-id/1258499?.
Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics, 13*(1), 27–35.
Grammig, J., Melvin, M., & Schlag, C. (2005). Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects. *Journal of Empirical Finance, 12*, 139–164.
Hagströmer, B., & Nordén, L. (2013). The diversity of high-frequency traders. *Journal of Financial Markets, 16*, 741–770.
Harris, F. H.d., McInish, T. H., & Wood, R. A. (2002). Security price adjustment across exchanges: An investigation of common factor components for Dow stocks. *Journal of Financial Markets, 5*, 277–308.
Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance, 50*(4), 1175–1199.
Hasbrouck, J. (2003). Intraday price formation in U.S. Equity index markets. *Journal of Finance, 58*(6), 2375–2399.
Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets, 16*, 649–679.
Hendershott, T., & Jones, C. M. (2005). Island goes dark: Transparency, fragmentation, and regulation. *Review of Financial Studies, 18*, 743–793.
Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *Journal of Finance, 66*(1), 1–32.
Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control, 12*, 231–254.
Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets, 3*, 205–258.
Pagnotta, E., & Philippon, T. (2012). *Competing on speed NBER working paper*. New York University: Stern School of Business.
Pascual, R., Pascual-Fuster, B., & Climent, F. (2006). Cross-listing, price discovery and the

informativeness of the trading process. *Journal of Financial Markets, 9*, 144–161.
Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics, 85*(4), 777–792.
Riordan, R., & Storkenmaier, A. (2012). Latency, liquidity and price discovery. *Journal of Financial Markets, 15*, 416–437.
Rogow, G. (2009, June 19). Rise of the (market) machines. *The Wall Street Journal*.
Skjeltorp, J. A., Sojli, E., & Tham, W. W. (2017). *Trading on algos. Working paper*. Norges Bank and Erasmus University of Rotterdam.
Stein, J. C. (2009). Presidential address: Sophisticated investors and market efficiency. *Journal of Finance, 64*(4), 1517–1548.
Stickel, S. E., & Verrecchia, R. E. (1994). Evidence that trading volume sustains stock price changes. *Financial Analysts Journal, 50*(6), 57–67.