How Smooth Is Price Discovery?

Evidence from Cross-listed Stock Trading

Haiqiang Chen[a], Paul Moon Sub Choi[b,*], Yongmiao Hong[a,c]

[a]*Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361005, China*

[b]*Ewha School of Business, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 120-750, Republic of Korea*

[c]*Department of Economics, Cornell University, 424 Uris Hall, Ithaca, New York 14853, U.S.A.*

This version: January 15, 2012

[*] Choi (Assistant Professor, corresponding author) can be contacted at paul.choi@ewha.ac.kr; College of Business Administration, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 120-750, Republic of Korea; tel.: +82-2-3277-3543; fax: +82-2-3277-2776. Chen (Assistant Professor) can be contacted at hc335@xmu.edu.cn; Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361005, China; tel.: +86-592-218-8827; fax: +86-592-218-7708. Hong (Professor) can be contacted at yh20@cornell.edu; Department of Economics, Cornell University, 424 Uris Hall, Ithaca, New York 14853, U.S.A.; tel.: +1-607-255-5130; fax: +1-607-255-2818. Standard disclaimer rules apply and all errors are of our own.

How Smooth Is Price Discovery? Evidence from Cross-listed Stock Trading

This version: January 15, 2012

Abstract

The adjustment to parity can be nonlinear for a cross-listed pair: Convergence may be quicker when the price deviation is sufficiently profitable. We propose a threshold error correction model (ECM) to gauge the market-respective information shares of Canadian listings traded on the Toronto Stock Exchange (TSX) and the New York Stock Exchange (NYSE). Since dynamics may alternatively be gradual, we further generalize the threshold framework to a smooth transition ECM. First, the TSX and the NYSE appear to have integrated over time. Second, parity-convergence accelerates upon discounts on the cross-listings on the NYSE. Third, we find a larger feedback from the NYSE if the price gap exceeds the threshold (required arbitrage return). Fourth, informed traders tend to cluster on the NYSE upon discounts on the cross-listings. Fifth, the information share and threshold are affected by the relative degree of private information, market friction and liquidity measures, and firm-level characteristics.

JEL Classification: C32; G15; G14

Keywords: Price discovery; Information share; Threshold error correction model; Smooth transition error correction model

1. Introduction

We contribute to the literature by implementing nonlinear error correction mechanisms in estimating the relative extent of exchange-respective contribution to price discovery (information share) of cross-listings and their original listings.[1] The existing methods assume linear convergence of price deviations to parity whereas we hinge our premise on the reality that (1) the premiums or discounts on cross-listings disappear quicker when they are profitably arbitrageable than otherwise; and (2) the convergence may occasionally be gradual and nonlinear.

Price discovery is the process by which information is priced in the market. When a security is traded in multiple markets, it is often of interest to determine where and how price discovery occurs. Harris et al. (1995) and Hasbrouck (1995) examine the information share of the NYSE of fragmented (multi-market traded) stocks on the NYSE and other U.S. exchanges, and confirm the leadership assumed by the NYSE. As for international cross-listing, Bacidore and Sofianos (2002) and Solnik et al. (1996) suggest that price discovery mostly takes place in the home market where substantial information originates. Eun and Sabherwal (2003) report the U.S. host exchanges determine the prices of Canadian cross-listings, however, to a lesser extent than the Toronto Stock Exchange (TSX) does.

In the literature, there are two broad approaches to estimating the contribution of each market to price discovery of fragmented listings. Hasbrouck's (1995) *innovation variance approach* extracts the information shares by employing variance decomposition based on the vector moving average representation of an error correction model (ECM). Harris et al.'s (1995, 2002) *common factor approach* employs permanent-transitory decomposition of a cointegrated system to estimate the information share of each market. As Eun and Sabherwal

[1] In this paper, we illustrate our methods using cross-listings. However, the methods can be applied to other informationlly linked markets, such as commodity futures, exchange rate, bond markets etc. See Liu and An (2011), Chen and Gau (2010), and Fricke and Menkhoff (2011).

(2003) point out, Hasbrouck's (1995) approach involves Cholesky factorization of the covariance matrix of the innovations to prices on various exchanges and yields multiple information shares. This may cause confounding identification of the venue of price discovery. Hasbrouck's (2002) modification can be numerically onerous in implementation.[2] In this paper, we expand Harris et al.'s (1995, 2002) platform and complement Hasbrouck's (1995) idea.

Harris et al. (1995) associate error correction dynamics with price discovery of cross-listed pairs which are cointegrated[3] by the law of one price. The cointegrating vectors of the vector ECM (VECM) represent the long-run equilibrium (near-parity condition), while the error correction terms characterize the convergence mechanism. Through representation, one can assess the relative extent of the contribution made by each market to price discovery of fragmented stocks using the estimates of adjustment coefficients. Harris et al. (2002) buttress the method earlier formulated in Harris et al. (1995) by incorporating a microstructure model where the price is assumed to be the sum of an efficient (permanent) price component and an error (transitory) term.[4]

However, an implicit assumption made by Harris et al.'s (1995, 2002) works is that adjustment to parity, the long-run equilibrium, is continuous and linear. Various economic circumstances challenge such restrictions, particularly where transaction costs and policy intervention are present. Given the complexity of trading rules and indirect transaction costs,

[2] See De Jong (2002), Harris et al. (2002), and Hasbrouck (2002) for further discussion.

[3] A group of multiple random-walk processes is cointegrated if, by definition, there exists a stationary linear combination of the processes. A time series is (weakly) stationary if the probability laws (of up to the second moments) are time-invariant.

[4] In Harris et al. (2002), the efficient price component is unobservable and reflects the underlying fundamentals. Gonzalo and Granger's (1995) permanent-transitory decomposition posits the permanent price as a linear combination of the observable prices where the normalized weights can be as market-respective information shares. The higher the normalized weight of an exchange, the bigger the influence on setting the permanent price. It can be shown that the normalized weights are orthogonal to the adjustment coefficient vector, which can be conveniently obtained from an ECM.

nonlinear convergence to parity captures the market to a higher proximity. The rationale of nonlinear modeling is straightforward. A relatively small deviation of the price of a cross-listing from its parity-implied price can be unarbitrageable if the dollar spread is insufficient to cover the fees, commissions, liquidity shortfalls, and other related costs. In this case, the dollar premium or discount behaves like a near-unit root process and will not converge to parity. Arbitrage forces will activate as the spread widens beyond the threshold, determined by transaction costs and associated risk premiums of arbitrage. There may be another case that there is more than a single break-even relative deviation in a cross-listed pair: The transition from a profitable status to an unarbitrageable one can be "smooth" due to multiple and overlapping regime-switching effects.

To date, we find a dearth of articles with a nonlinear framework in the literature. Among them, Rabinovitch et al. (2003) use a nonlinear threshold model to estimate the adjustment dynamics of the return deviations for 20 Chilean and Argentine cross-listings. Koumkwa and Susmel (2008) suggest two nonlinear adjustment models: The exponential smooth transition autoregressive (ESTAR) and the logarithmic smooth transition autoregressive (LSTAR) to delineate the relative premiums[5] of Mexican ADRs. Chung et al. (2005) study the dynamic relationship between the prices of three Taiwanese ADRs and their underlying stocks using a threshold VECM.[6] These mentioned articles are devoted to the asset pricing aspect of cross-listings, whereas our novel approaches focus on the price discovery of muti-market listings using nonlinear frameworks.

We extend Harris et al.'s (1995, 2002) ECM to estimate the exchange-respective

---

[5] We define the "relative premium" as the percentage premium of a cross-listed stock traded on a foreign exchange against the home market share, adjusted by the exchange rate. The term "cross-listing premium" defined by Doidge et al. (2004) is the excess value of foreign firms cross-listed in the U.S. relative to those not in terms of Tobin's (1969) *q* ratio.

[6] As a linear modeling precursor, see Kim et al. (2000) for vector autoregressive (VAR) and seemingly unrelated estimation (SURE) methods that analyze adjustments in ADR-implied prices.

information shares of Canadian cross-listed pairs traded on the NYSE and the TSX[7] by, firstly, considering threshold cointegration per Balke and Fomby (1997) and by, secondly, generalizing it to a smooth transition version. Departing from linear modeling, our information share measures are derived from (1) the outer-regime adjustment coefficients based on a two-regime threshold ECM; and (2) the average coefficient estimates based on a smooth transition ECM, respectively. The former method is intuitively appealing whereas the latter amendment risks less model misspecification.

Our alternative methods have many advantages. To list a few of them, first, we theoretically depict and empirically analyze the discrete dynamics of both abrupt and smooth parity-convergences, which are frequently observed in the market due to various risk factors such as information asymmetry and market friction. Second, a large deviation from parity far beyond the threshold (extreme regime), e.g. a very profitable arbitrage opportunity, is more likely to reflect information shocks than a small deviation which can be due to noise trading.[8] Third, a gradual and nonlinear conversion to parity can be better detected with our smooth transition ECM framework. Thus, we believe our methods can capture relative contribution to price discovery to a higher degree compared to the existing linear approaches in the literature, which circumvent the time and regime-contingent characteristics of information shocks.

---

[7] We choose to study Canadian stocks listed in the U.S. for several reasons following Eun and Sabherwal (2003) and Chen and Choi (2012). First, Canadian equities are the largest group of stocks cross- listed in the U.S. from a single country. Second, many of these Canadian stocks trade actively on both the NYSE and the TSX which is essential for conducting intraday tests. Third, the trading hours of the TSX coincide with that of the NYSE (9:30AM— 4:00PM, EST), Finally, Canadian stocks trade in the U.S. as ordinary shares due to compatible accounting standards, whereas most other cross- listed shares are ADRs issued by U.S. custodian banks.

[8] A similar idea is illustrated by Gonzalo and Marinz (2006) in a model of price discovery for stocks traded in a single market. In their model, only new information which implies a profit greater than the transaction cost, measured by bid-ask spread, will be translated into the transaction price. In other words, the shocks that drive the efficient price component must be “big” shocks to the transaction price. The transactions of uninformed agents cannot generate big inefficient changes in the transaction prices, because informed traders will arbitrate the situation. Therefore, the shocks driving the pricing error component by uninformed traders must be “small” shocks to the transaction price.

In addition, we identify and explicate the factors that affect the estimated information shares and thresholds. Specifically, our empirical findings can be summarized into fivefold: First, parity-convergence accelerates upon discounts on the cross-listings on the NYSE. Second, we find a larger feedback from the NYSE if the price gap exceeds the threshold. Third, informed traders tend to cluster on the NYSE upon discounts on the cross-listings. Forth, the information share and threshold estimates are affected by the relative degree of private information, market friction and liquidity measures, and firm-level characteristics.

The remainder of this paper is organized as follows. We provide a price discovery model of cross-listings to illustrate the role of arbitrageurs in Section 2. Section 3 develops the econometric models implied from the price discovery model developed in Section 2: The existing standard ECM and our threshold and smoon transition ECMs. Section 4 describes the data and presents statistical test results. Discussion of the estimation and panel regression analyses are shown in Section 5. We conclude in Section 6.

2. Price discovery of cross-listings

In this section, we develop an equilibrium model to characterize the interactive dynamics of a cross-listed pair simutaneously traded on two separate exchanges. Arbitrageurs linking the two markets may be subject to market frictions, such as transaction fees, capital constraints etc. Throughout the model, we emphasize the role of arbitragers in the process of inter-market price discovery.

We first assume that there are two cross-border stock exchanges: The TSX and the NYSE. We conveniently index the respective markets: $i = 1, 2$. We further assume that there are N_1 investors who only trade at home (TSX) and N_2 only abroad (NYSE), and N_3 arbitrageurs who interplay in both markets. We assume choice of exchanges by the former two trader types (one-market traders) is exogeneous and is due to various reasons such as distance, language, institutional constraints, transaction costs, etc.

We subsequently specify the behavior of one-market trader in market i. At time t, for trader j, we let E_{jt}^{i} be her endowment and μ_{jt}^{i} be the reservation price of the listing in market i. Given an available U.S. dollar-translated market price p_{it}, her demand function can be conjectured as

$$X_{ijt} = E_{jt}^{i} - \beta\left(\mu_{jt}^{i} - p_{it}\right) \quad \text{for } j = 1,2,\cdots,N_i, \tag{1}$$

where $\beta > 0$ is the demand elasticity, same for all one-market traders.

We now consider the demand function of arbitrageurs. Arbitrageurs are initially endowed with no seed money. Arbitrageurs "buy low and sell high" between the two markets, thus their demand function only depends on the cross-border price deviation. Given respective prices p_{1t} and p_{2t}, arbitrageur j would submit her buy order in market 1 as

$$X_{1jt}^{A} = \beta_{jt}^{A}(p_{2t} - p_{1t}) \quad \text{for } j = 1,2,\cdots,N_3, \tag{2}$$

where $\beta_{jt}^{A} > 0$ is the demand elasticity.[9] Since she hedges perfectly,

$$X_{2jt}^{A} = -X_{1jt}^{A} = \beta_{jt}^{A}(p_{1t} - p_{2t}), \tag{3}$$

i.e. her short position in one market always equals the long position in the other market.

In equilibrium, the two exchanges clear as

$$\sum_{j=1}^{N_1} E_{jt}^{1} = \sum_{j=1}^{N_1} X_{1jt} + \sum_{j=1}^{N_3} \beta_{jt}^{A}(p_{2t} - p_{1t}), \tag{4}$$

$$\sum_{j=1}^{N_2} E_{jt}^{2} = \sum_{j=1}^{N_2} X_{2jt} + \sum_{j=1}^{N_3} \beta_{jt}^{A}(p_{1t} - p_{2t}). \tag{5}$$

Solving the market clearing conditions for equilibrium prices of the cross-listed pair yields

$$p_{1t} = \frac{\left(\beta N_2 + \sum_{j=1}^{N_3} \beta_{jt}^{A}\right) N_1 \mu_t^1 + \sum_{j=1}^{N_3} \beta_{jt}^{A} N_2 \mu_t^2}{\left(\beta N_2 + \sum_{j=1}^{N_3} \beta_{jt}^{A}\right) N_1 + \sum_{j=1}^{N_3} \beta_{jt}^{A} N_2}, \tag{6}$$

[9] Following Garbade and Silber (1983), the demand elasticity for arbitrageurs ($\beta_{j,t}^{A}$) is assumed to be finite due to market friction.

$$p_{2t} = \frac{\left(\beta N_1+\sum_{j=1}^{N_3}\beta_{jt}^A\right)N_2\mu_t^2+\sum_{j=1}^{N_3}\beta_{jt}^A N_1\mu_t^1}{\left(\beta N_1+\sum_{j=1}^{N_3}\beta_{jt}^A\right)N_2+\sum_{j=1}^{N_3}\beta_{jt}^A N_1}, \tag{7}$$

where $\mu_t^1 = \frac{1}{N_1}\sum_{j=1}^{N_1}\mu_{jt}^1$ and $\mu_t^2 = \frac{1}{N_2}\sum_{j=1}^{N_2}\mu_{jt}^2$ are market average reservation prices.

In order to derive dynamic price relationships, we further specify an evolution mechanism of the reservation prices μ_{jt}^1 and μ_{jt}^2, following Garbade and Silber (1983), as

$$\mu_{jt}^i = p_{it-1} + v_t + \varepsilon_{jt}^i \quad \text{for } i = 1,2,\ j = 1,2,\cdots,N_i. \tag{8}$$

As market i clears at the end of period $t-1$ with a partial equilibrium price p_{it-1}, each trader decides to hold her share of asset towards her endowment in the subsequent period t, E_{jt}^i. As new information on the issuer, v_t, common to all investors in both markets arrives, the trader formulates her new reservation prices μ_{jt}^i with an idiosyncratic error ε_{jt}^i. We assume v_t and ε_{jt}^i are i.i.d normal random variables with mean zero and constant variance, respectively.

In aggregate, the market reservation prices μ_t^1 and μ_t^2 can be expressed as

$$\mu_t^1 = \frac{1}{N_1}\sum_{j=1}^{N_1}\mu_{jt}^1 = p_{1t-1} + v_t + \frac{1}{N_1}\sum_{j=1}^{N_1}\varepsilon_{jt}^1, \tag{9}$$

$$\mu_t^2 = \frac{1}{N_2}\sum_{j=1}^{N_2}\mu_{jt}^2 = p_{2t-1} + v_t + \frac{1}{N_2}\sum_{j=1}^{N_2}\varepsilon_{jt}^2. \tag{10}$$

Plugging μ_t^1 and μ_t^2 into the equations (6) and (7), we have,

$$p_{1t} = \frac{(\beta N_2+\sum_{j=1}^{N_3}\beta_{jt}^A)N_1 p_{1t-1}+\sum_{j=1}^{N_3}\beta_{jt}^A N_2 p_{2t-1}}{(\beta N_2+\sum_{j=1}^{N_3}\beta_{jt}^A)N_1+\sum_{j=1}^{N_3}\beta_{jt}^A N_2} + v_t + \tilde{\varepsilon}_t^1, \tag{11}$$

$$p_{2t} = \frac{(\beta N_1+\beta^A N_3)N_2 p_{2t-1}+\sum_{j=1}^{N_3}\beta_{jt}^A N_1 p_{1t-1}}{(\beta N_1+\sum_{j=1}^{N_3}\beta_{jt}^A)N_2+\sum_{j=1}^{N_3}\beta_{jt}^A N_1} + v_t + \tilde{\varepsilon}_t^2, \tag{12}$$

where

$$\tilde{\varepsilon}_t^1 = \frac{(\beta N_2+\sum_{j=1}^{N_3}\beta_{jt}^A)\sum_{j=1}^{N_1}\varepsilon_{jt}^1+\sum_{j=1}^{N_3}\beta_{jt}^A\sum_{j=1}^{N_2}\varepsilon_{jt}^2}{(\beta N_2+\sum_{j=1}^{N_3}\beta_{jt}^A)N_1+\sum_{j=1}^{N_3}\beta_{jt}^A N_2}, \tag{13}$$

$$\tilde{\varepsilon}_t^2 = \frac{(\beta N_1+\sum_{j=1}^{N_3}\beta_{jt}^A)\sum_{j=1}^{N_2}\varepsilon_{jt}^2+\sum_{j=1}^{N_3}\beta_{jt}^A\sum_{j=1}^{N_1}\varepsilon_{jt}^1}{(\beta N_1+\sum_{j=1}^{N_3}\beta_{jt}^A)N_2+\sum_{j=1}^{N_3}\beta_{jt}^A N_1}. \tag{14}$$

An equivalent matrix representation prescribes

$$\begin{pmatrix} p_{1t} \\ p_{2t} \end{pmatrix} = \begin{pmatrix} 1-a_t, a_t \\ b_t, 1-b_t \end{pmatrix} \begin{pmatrix} p_{1t-1} \\ p_{2t-1} \end{pmatrix} - \begin{pmatrix} v_t + \tilde{\varepsilon}_t^1 \\ v_t + \tilde{\varepsilon}_t^2 \end{pmatrix} \tag{15}$$

where

$$a_t = \frac{\sum_{j=1}^{N_3} \beta_{jt}^A N_2}{\beta N_2 N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_2}, \tag{16}$$

$$b_t = \frac{\sum_{j=1}^{N_3} \beta_{jt}^A N_1}{\beta N_2 N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_2}. \tag{17}$$

We can obtain the following bivariate VECM by subtracting $(p_{1t-1}, p_{2t-1})^T$ from both sides:

$$\begin{pmatrix} \Delta p_{1t} \\ \Delta p_{2t} \end{pmatrix} = \begin{pmatrix} -a_t, a_t \\ b_t, -b_t \end{pmatrix} \begin{pmatrix} p_{1t-1} \\ p_{2t-1} \end{pmatrix} - \begin{pmatrix} v_t + \tilde{\varepsilon}_t^1 \\ v_t + \tilde{\varepsilon}_t^2 \end{pmatrix}. \tag{18}$$

The above VECM describe the short term dynamics toward the long run equilibrium, given the cointegrating vector $(1, -1)^T$. The short term adjustment coefficients a_t and b_t for respective prices, p_{1t} and p_{2t}, reflect their responses to deviations from the long run equilibrium in respective markets. We can apply the permanent transitory decomposition (Granger and Gonzalo, 1995) to the above VECM: The permanent component is a linear combination of $(p_{1t},\ p_{2t})$ formed by the scaled orthogonal vector of the adjustment coefficient vector (a_t, b_t). Specifically, the permanent component is given by

$$f_t = \frac{b_t}{a_t+b_t} p_{1t} + \frac{a_t}{a_t+b_t} p_{2t}. \tag{19}$$

where

$$\frac{b_t}{a_t+b_t} = \frac{N_1}{N_1+N_2}, \tag{20}$$

$$\frac{a_t}{a_t+b_t} = \frac{N_2}{N_1+N_2}. \tag{21}$$

$\frac{b_t}{a_t+b_t}$ and $\frac{a_t}{a_t+b_t}$ capture the contribution share of each price to the permanent component: They reflect the respective *information shares* of markets 1 and 2 towards determining the long run equilibrium price. In other words, they are relative measures of market specific contribution to price discovery of the cross-listed pair.

Define $\Delta p_t \equiv p_{2t} - p_{1t}$ as the dollar premium on the cross-listing against its original listing. It can be shown that

$$\Delta p_t = \delta_t \Delta p_{t-1} + e_t, \tag{22}$$

where

$$\begin{aligned}\delta_t &= 1 - a_t - b_t \\ &= 1 - \frac{\sum_{j=1}^{N_3} \beta_{jt}^A N_2}{\beta N_2 N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_2} - \frac{\sum_{j=1}^{N_3} \beta_{jt}^A N_1}{\beta N_2 N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_2} \\ &= \frac{\beta N_2 N_1}{\beta N_2 N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_1 + \sum_{j=1}^{N_3} \beta_{jt}^A N_2}.\end{aligned} \tag{23}$$

Following Garbade and Silber (1983), δ_t measures the reciprocal convergence speed of the two market prices to their long run equilibrium: Convergence is quicker the smaller δ_t is.

3. Error correction models

The equilibrium model constructed in Section 2 characeterizes measures of market-wise contribution to price discovery, which is related to the relative population of respective market participants (Equations (20) and (21)). However, it poses an empirical challenge since the "headcounts" are usually unknown. Fortunately, we can estimate the adjustment coefficients a_t and b_t via the error correction model (Equation (18)) which only contains the information of market prices. Another hurdle is that the adjustment coefficients a_t and b_t are time-varying. In order to estimate the model, additional restrictions are necessary to characterize the time paths of a_t and b_t. In the following three subsections, by adopting different assumptions on arbitraguers, the equilibrium model presented in Section 2 can generate three different version of ECMs: Standard ECM, threshold ECM, and smooth transition ECM. These models provide various estimates of the adjustment coefficients and, thus, information shares.

3.1. Standard error correction model

We begin with a standard ECM, where the adjustment coeffcients, a_t and b_t, are assumed to be constant in Equation (18). All arbitrageurs are homogeneous and the market is perfectly competitive, i.e., there are neither transaction costs nor other market frictions. Under these assumptions, we have $\beta_{jt}^A = \beta^A$ for all j and t, with $\beta^A > 0$. It follows that

$$a_t = \frac{N_3\beta^A N_2}{\beta N_2 N_1 + N_3\beta^A N_1 + N_3\beta^A N_2} \equiv a, \tag{24}$$

$$b_t = \frac{N_3\beta^A N_1}{\beta N_2 N_1 + N_3\beta^A N_1 + N_3\beta^A N_2} \equiv b, \tag{25}$$

which are constant for all t, and

$$\begin{pmatrix}\Delta p_{1t}\\ \Delta p_{2t}\end{pmatrix} = \begin{pmatrix}-a, a\\ b, -b\end{pmatrix}\begin{pmatrix}p_{1t-1}\\ p_{2t-1}\end{pmatrix} - \begin{pmatrix}v_t + \tilde{\varepsilon}_t^1\\ v_t + \tilde{\varepsilon}_t^2\end{pmatrix}$$

$$= \begin{pmatrix}a\\ -b\end{pmatrix}(p_{1t-1} - p_{2t-1}) - \begin{pmatrix}v_t + \tilde{\varepsilon}_t^1\\ v_t + \tilde{\varepsilon}_t^2\end{pmatrix}. \tag{26}$$

Define the dollar premium on the cross-listing against its original listing as

$$\kappa_t \equiv p_{2t} - p_{1t}. \tag{27}$$

A standard ECM for the bivariate cointegrated system of the cross-listed pair can be structured as

$$\Delta p_{1t} = \beta_{10} + \alpha_1\kappa_{t-1} + \sum_{j=1}^{m_1}\beta_{1j}\Delta p_{1t-j} + \sum_{j=1}^{m_2}\tilde{\beta}_{1j}\Delta p_{2t-j}, \tag{28}$$

$$\Delta p_{2t} = \beta_{20} + \alpha_2\kappa_{t-1} + \sum_{j=1}^{m_1}\beta_{2j}\Delta p_{1t-j} + \sum_{j=1}^{m_2}\tilde{\beta}_{2j}\Delta p_{2t-j}, \tag{29}$$

where κ_{t-1} gives the remaining cross-listing dollar premium or cointegrating residual. α_1 and α_2 are the adjustment coefficients of the TSX and the NYSE, respectively: They describe how much deviation will be subsequently adjusted to restore the long run equilibrium in each series. Per *Granger Representation Theorem* (Engle and Granger, 1987; Engle and Yoo, 1987), if p_{1t} and p_{2t} are cointegrated, then at least one of α_1 and α_2

must be nonzero. In other words, p_{1t} and p_{2t} or both, will adjust fractionally to restore parity in the long run.

Harris et al. (1995, 2000) propose to use the ECM adjustment coefficients to estimate the relative extent of exchange-respective contribution to price discovery (information share) of shares whose order purchases are fragmented across multiple markets. For a Canadian company originally listed on the TSX and cross-listed on the NYSE, the proportion of the adjustments that takes place on the TSX out of the total adjustments occurred on both exchanges is the share of the home exchange in contribution to setting the long-run equilibrium price as a result of synchronous cross-border stock trading. In an extreme case where there is no feedback from the NYSE such that $\alpha_2 = 0$, then the NYSE has no contribution to price discovery of the cross-listed pair. Eun and Sabherwal (2003) further define the respective information shares of the TSX and the NYSE as

$$\mathrm{IS}_1 \equiv \frac{|\alpha_1|}{|\alpha_1|+|\alpha_2|} \text{ and } \mathrm{IS}_2 \equiv \frac{|\alpha_2|}{|\alpha_1|+|\alpha_2|}. \tag{30}$$

Suppose $p_{1t-1} < p_{2t-1}$ in the previous period $(t-1)$, then a likely scenario to reduce the gap between the two prices is: p_{1t} increases or p_{2t} to decreases, or both. In this case one can conjecture that α_1 is non-positive and α_2 is non-negative. There two other possibilities: (1) p_{1t-1} decreases but p_{2t-1} decreases more; or (2) p_{1t-1} increases but p_{2t-1} increases less.[10] As Eun and Sabherwal (2003) put the latter two turnouts are very unlikely, they are not considered in our study. One can analogously design a similar adjustment mechanism to show that α_1 is non-positive and α_2 is non-negative for the symmetric situation when $p_{1t-1} > p_{2t-1}$. Therefore, we define the exchange-respective information shares of the TSX and the NYSE as

$$\mathrm{IS}_1 \equiv \frac{-\alpha_1}{-\alpha_1+\alpha_2} \text{ and } \mathrm{IS}_2 \equiv \frac{\alpha_2}{-\alpha_1+\alpha_2}. \tag{31}$$

[10] These odds may reflect the under-reaction to the information share of the market. When information incorporation takes multiple periods, the price adjustment should persist unilaterally during then.

3.2. Threshold error correction model

In reality, the market is imperfect due to various sources of market friction such as transaction costs, direct and indirect trading barriers, etc. We let the threshold (γ) measure the sum of transaction costs and risk premiums required from arbitrageurs (required arbitrage return). Arbitrage opportunities exist when

$$\kappa_t \equiv p_{2t} - p_{1t} < -\gamma \text{ or } \kappa_t > \gamma, \tag{32}$$

which becomes $|\kappa_t| > \gamma$.[11] We still assume all arbitragers are homogeneous, such that they share the common demand elasticity. Under these assumptions, we have

$$\beta_{jt}^{A} = \begin{Bmatrix} 0 & \text{if } |\kappa_t| \leq \gamma \\ \beta^{A} > 0 & o.w. \end{Bmatrix}. \tag{33}$$

Now, cointegration between p_{1t} and p_{2t} is dormant within a range of disequilibrium but the error correction dynamics becomes active once the cross-listing dollar premium sufficiently digresses from parity beyond the threshold. Balke and Fomby (1997) propose this regime-switching mechanism as *threshold cointegration*, and the implied error correction dynamics can be characterized by a threshold ECM, given by

$$\Delta p_{1t} = \begin{Bmatrix} \beta_{110} + \alpha_{11}\kappa_{t-1} + \sum_{j=1}^{m_1} \beta_{11j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{11j}\Delta p_{2t-j}, \text{ if } |\kappa_{t-1}| \leq \gamma \\ \\ \beta_{120} + \alpha_{12}\kappa_{t-1} + \sum_{j=1}^{m_1} \beta_{12j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{12j}\Delta p_{2t-j}, \text{ if } |\kappa_{t-1}| > \gamma \end{Bmatrix}, \tag{34}$$

$$\Delta p_{2t} = \begin{Bmatrix} \beta_{210} + \alpha_{21}\kappa_{it-1} + \sum_{j=1}^{m_1} \beta_{21j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{21j}\Delta p_{2t-j}, \text{ if } |\kappa_{t-1}| \leq \gamma \\ \\ \beta_{220} + \alpha_{22}\kappa_{it-1} + \sum_{j=1}^{m_1} \beta_{22j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{22j}\Delta p_{2t-j}, \text{ if } |\kappa_{t-1}| > \gamma \end{Bmatrix}. \tag{35}$$

In the middle regime when $|\kappa_{t-1}| \leq \gamma$, there are neither market forces nor arbitrageurs to sustain cointegration of the pair of prices. In other words, unless the pair shows a

[11] Transaction costs of cross-border arbitrage are comprised of the bid-ask spreads of the prices on both exchanges and the foreign exchange rate, fixed costs, and liquidity shortfalls. Chen and Choi (2012) find the relative premium of a Canadian cross-listing on the NYSE, on average, includes an adverse-selection risk premium due cross-border imbalance in private information on the issuing firm. Along with the asymmetric information component, macroeconomic factors, such as GDP growth rates and interest rates, may also affect determining the threshold.

significant price gap exceeding the threshold minimum profit, the adjustment coefficients are zeroes ($\alpha_{11} = \alpha_{21} = 0$) and, thus, neither price (p_{1t} nor p_{2t}) appropriately reflects risks. We define the information share, or the relative measure of contribution to price discovery, for respective market using the outer regime coefficient estimates[12] (α_{12} and α_{22}):

$$\mathrm{IS}_1 \equiv \frac{|\alpha_{12}|}{\alpha_{12}+|\alpha_{12}|} \text{ and } \mathrm{IS}_2 \equiv \frac{\alpha_{22}}{|\alpha_1|+\alpha_{12}}. \tag{36}$$

3.3. Smooth transition error correction model

A common assumption for the standard and threshold ECMs is homogeneity of arbitrageurs. In this subsection, we relax this assumption to allow for heterogeneous arbitrageurs since they may require differing thresholds (γ_j's) to establish their respective positions. For example, using a database of arbitrage transactions on the NYSE, Neal (1996) finds that larger mispricings are positively related to the frequency of arbitrage porfolio construction. Moreover, fees paid by institutional investors depend on the arrangement between the investors and their executing brokers. The opportunity cost faced by capital-constrained arbitrageurs can be another reason for different threshold values: The investors with stricter capital constraints will skip small mispricings to wait for larger ones. In sum, there does not exist a unique break-even point for all arbitraguers. Specifically, we assume, for arbitrageur $j = 1,2,\cdots,N_3$,

$$\beta_{jt}^A = \begin{Bmatrix} 0 & \text{if } |\kappa_t| \le \gamma_j \\ \beta^A > 0 & o.w. \end{Bmatrix}. \tag{37}$$

The "aggregate" threshold is a smoothing function of price deviations such that

$$\sum_{j=1}^{N_3} \beta_{jt}^A = N_3 E(\beta_j^A) = N_3 \left\{ \int_{-\infty}^{|\kappa_{t-1}|} \beta^A dF(\gamma) + \int_{|\kappa_{t-1}|}^{\infty} \beta^A dF(\gamma) \right\} \equiv \beta^A(\kappa_{t-1}), \tag{38}$$

where $F(\gamma)$ is the cumulative density function of γ_j across all j. Further

$$a_t = \frac{\beta^A(|\kappa_{t-1}|)N_2}{\beta N_2 N_1 + \beta^A(|\kappa_{t-1}|)N_1 + \beta^A(|\kappa_{t-1}|)N_2} \equiv \alpha_1(\kappa_{t-1}), \tag{39}$$

[12] Eun and Sabherwal (2003) estimate the adjustment coefficients in every period using a linear ECM following Harris et al. (1995).

$$b_t = \frac{\beta^A(|\kappa_t|)N_1}{\beta N_2 N_1 + \beta^A(|\kappa_t|)N_1 + \beta^A(|\kappa_t|)N_2} \equiv \alpha_2(\kappa_{t-1}). \tag{40}$$

By plugging a_t and b_t into Equation (15), we obtain a smooth transition ECM:

$$\Delta p_{1t} = \beta_{10} + \alpha_1(|\kappa_{t-1}|)\kappa_{it-1} + \sum_{j=1}^{m_1} \beta_{1j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{1j}\Delta p_{2t-j}, \tag{41}$$

$$\Delta p_{2t} = \beta_{20} + \alpha_2(|\kappa_{t-1}|)\kappa_{it-1} + \sum_{j=1}^{m_1} \beta_{2j}\Delta p_{1t-j} + \sum_{j=1}^{m_2} \tilde{\beta}_{2j}\Delta p_{2t-j}. \tag{42}$$

The "average" information shares of respective market can be defined as

$$IS_1 \equiv \frac{|\mathbb{E}(\alpha_1(\kappa_{t-1}))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1}))| + |\mathbb{E}(\alpha_2(\kappa_{t-1}))|}, \tag{43}$$

$$IS_2 \equiv \frac{|\mathbb{E}(\alpha_2(\kappa_{t-1}))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1}))| + |\mathbb{E}(\alpha_2(\kappa_{t-1}))|}, \tag{44}$$

where

$$\mathbb{E}(\alpha_1(\kappa_{t-1})) = \frac{1}{T}\sum_{t=1}^{T} \alpha_1(\kappa_{t-1}), \tag{45}$$

$$\mathbb{E}(\alpha_2(\kappa_{t-1})) = \frac{1}{T}\sum_{t=1}^{T} \alpha_2(\kappa_{t-1}). \tag{46}$$

In order to see whether the informed traders will choose the market with a relative discount, the "cross" information share of an exchange can be defined for cases with a discount or premium on the cross-border trading venue:

$$IS_{11} \equiv \frac{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}>0))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}>0))| + |\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}>0))|}, \tag{47}$$

$$IS_{21} \equiv \frac{|\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}>0))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}>0))| + |\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}>0))|}, \tag{48}$$

$$IS_{12} \equiv \frac{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}<0))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}<0))| + |\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}<0))|}, \tag{49}$$

$$IS_{22} \equiv \frac{|\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}<0))|}{|\mathbb{E}(\alpha_1(\kappa_{t-1})I(\kappa_{t-1}<0))| + |\mathbb{E}(\alpha_2(\kappa_{t-1})I(\kappa_{t-1}<0))|}. \tag{50}$$

For a cross-listed pair, IS_{11} and IS_{21} are the information shares of the TSX and the NYSE, respectively, when the TSX-listing trades at a discount ($\kappa_{t-1} > 0$); IS_{12} and IS_{22} when the cross-listing posts a discount ($\kappa_{t-1} < 0$).

4. Data

56 TSX-NYSE pairs are identified through the sample period: January 1, 1998, through December 31, 2000. In order to estimate asymmetric-information and market-friction measures, high-frequency data are required for the shares co-listed on the TSX and the NYSE, and the U.S.-Canada exchange rate. Accordingly, the tick-by-tick trade and quote data for the TSX-listed Canadian stocks and the Trade-And-Quote (TAQ) data of their cross-listings on the NYSE through the period are used. The exchange rate intraday data was purchased from Olson & Associates. The co-listings on the TSX and the NYSE trade from 9:30AM to 4:00PM, EST, and the price deviations are calculated with simultaneous observations. On an average trading day, we observe 40 data points for each pair. Our sample period covers 772 trading days, but not all stocks are in pairs throughout the sample period. We require that for each firm-year the prices be observed in the two markets for at least 6 consecutive months. Further, having dropped thinly traded stocks on both exchanges, our final sample includes 40 cross-listed pairs and 104 firm-years.

4.1. Cointegration analysis

We first examine whether pairs of times series on the TSX and the NYSE price series are unit roots or not. We use the augmented Dickey and Fuller's (1981) (ADF) test, which considers lagged first differences of time series in the specification. If the test statistic is too large, then we reject the null hypothesis of unit root and conclude that the time series is stationary. As a result, the null hypothesis was rejected only for four out of 104 firm-years, at a five percent significance level. Thus, we conclude that both price series in our sample are, overall, first-order integrated ($I(1)$) or unit units.

We subsequently examined, using Johansen's (1991) test, cointegration between the two price series. The S&P TSX Composite and the S&P 500 indices (market indices of the TSX and the NYSE, respectively) were not included in the cointegration system since Eun and Sabherwal (2003) find that the estimated coefficients of the two index series are statistically insignificant. Since we have two price series in each regression equation, there is at most one cointegrating vector. We estimated the cointegrating vector ($(1,-b)^T$) for each cross-listed pair in each year. Our results show that most of the estimated cointegrating vectors are $(1,-1)^T$, which is the expected values according to the law of one price. Table 1 reports the summary statistics of the normalized estimation of the cointegrating vector for p_{1t} and p_{2t}, and the *t*-statistics for the null hypothesis attest that the cointegrating vector equals $(1,-1)^T$.

[Insert Table 1 about here.]

In Table 1, we see that the median of the normalized estimates throughout the sample is $(1,-1)^T$ which confirms that the Canadian cross-listed pairs tend to follow the law of one price and are, therefore, cointegrated. Given the estimated cointegrating vector $(1,-b)^T$, the estimated cross-listing dollar premium is $\kappa_t \equiv p_{2t} - p_{1t}$. We, then, test κ_t's for stationarity per the ADF test and find that only 3 out of 104 firm-years do not reject the null hypothesis of unit root. Thus, we conclude that the TSX-NYSE cross-listed pairs are cointegrated.

4.2. Nonlinearity tests

The law of one price suggests that two market prices for the same stock should not drift far from each other. This relationship is confirmed by the cointegration analysis in the previous section. However, linear adjustment dynamics is not necessarily prescribed by market efficiency assumptions. Given various market frictions, such as transactions costs and short sale limitations, thus, it is more likely that a nonlinear model, such as a threshold or smooth transition cointegration model, can provide a better description of the convergence procedure between two market prices. In this section, we conduct several nonlinearity tests in the course of short-run adjustment dynamics to long-run parity equilibrium.

We estimate the aforementioned symmetric bivariate threshold ECM model (Subsection 3.2) and apply the supremum-Lagrangian multiplier (supLM) test to check the nonlinearity. This test also has power to detect smooth transition error correction dynamics per Hansen and Seo (2002). We use Akaike's (1974) and Schwart's (1978) Bayesian information criteria to choose the number of lags, and consistently choose the lag length of 1 ($m_1 = m_2 = 1$). The cointegrating vector is given as $(1, -1)^T$, following the results of cointegration tests. The model is estimated by the maximum likelihood method described in Appendix A. The model is estimated in each year for each pair and the results are reported in Table 2.

[Insert Table 2 about here.]

Panel A of Table 2 exhibits the summary statistics of the threshold estimates and test statistics. The *p*-values are computed by the parametric bootstrap method suggested by Hansen and Seo (2002). We find that the mean and median of supLMs of the whole sample are 73.52 and 31.05, which exceeds 95% critical values 24.78 and 23.10. Therefore, on average, we can reject the null hypothesis of no threshold effect.

To further confirm the test results, we apply a combined *p*-value test on all firm-years. Let p_i be the asymptotic *p*-value of supLM test for each individual stock-year i, for $i = 1$, $2, \ldots, N$, where N is the total number of firm-years. We combine all *p*-values (p_i's) to construct the *Z*-test statistic proposed by Choi (2001):

$$Z \equiv \frac{1}{2\sqrt{N}} \sum_{i=1}^{N} \{-2\ln(p_i) - 2\}, \tag{51}$$

which is asymptotically standard normal under the null hypothesis of no threshold effect. In our untabulated case, the combined *p*-value test statistic Z is 33.41, significantly rejecting the null hypothesis with a 5% critical value at 1.96. Overall, we conclude that there exists nonlinearity in the course of parity convergence.[13]

It may be interesting to examine whether the threshold effect takes place on the

[13] We also repeated the estimation and testing procedures when the cointegrating vector is estimated from the data rather than restricted to $(1, -1)^T$. The results are qualitatively equivalent and the conclusion remains unaffected.

coefficients of error correction terms or short-term dynamics terms in the threshold ECM of Equations (34) and (35). We separately test the threshold effects on these coefficients. Panel B of Table 2 report the test results. The first two columns ($\text{Wald}_{\text{ECM1}}$) report the Wald statistics for testing the null hypothesis on the error correction terms: H_0: $\alpha_{11} = \alpha_{12}$. In other words, we test whether the adjustment coefficients associated with prices of TSX-listings (p_{1t}'s) are different within and beyond the thresholds. Likewise, the third and forth columns ($\text{Wald}_{\text{ECM2}}$) are testing the same hypothesis for the NYSE-listings: H_0: $\alpha_{21} = \alpha_{22}$. The last four columns (Wald_{DC1}, Wald_{DC2}) report the Wald statistics for the following null hypotheses of no threshold effect: H_0: $\beta_{111} = \beta_{121}$, $\tilde{\beta}_{111} = \tilde{\beta}_{121}$; and H_0: $\beta_{211} = \beta_{221}$, $\tilde{\beta}_{211} = \tilde{\beta}_{221}$. Rejecting these null hypotheses will lead to accepting the threshold effect on the short-term dynamics terms in the threshold ECM of the TSX and the NYSE, respectively. The untabulated, *Z*-test statistics (Choi, 2011) of the null hypotheses of error correction coefficients are 16.80 and 19.68, respectively, while those of short-term dynamics coefficients are 13.95 and 14.44, respectively. Thus, for both exchanges, we conclude that (1) there is nonlinearity in both error correction and short-term dynamics terms; and (2) the threshold effect is more likely to take place in the error correction terms. Put another way, (1) nonlinearity in the parity-convergence of a Canadian cross-listed pair is more persistent than transitory; and (2) the mean-reversion of a diverged pair tend to take place over a "brief" matter of time.

5. Empirics

5.1. Estimation

5.1.1. Microstructure measures

Unlike the NYSE, which is a specialist-based auction exchange, the TSX is an electronic exchange, which uses a Central Limit Order Book (CLOB) system, where orders are required to be posted in the book to be valid.[14] By studying decrements in the inside depth on one side of the quote that correspond to uncommon trade sizes (such as a trade of 1,300 shares), matching trades with prevailing quotes with a five-second lead (Lee and Ready, 1991) is reasonable: A trade is considered buyer-initiated if it is higher than the five-second earlier mid-quote, and seller-initiated if lower.[15] We construct datasets for estimation of the PIN following Easley et al. (1996, 2002). The NYSE-resident specialists are central to the theory of the PIN (Easley et al., 2001; Duarte and Young, 2008). There are official market makers, known as registered traders, on the TSX whose function is akin to that of the NYSE specialists. Thus, a comparison of trade informedness on the two exchanges by the PIN is deemed appropriate.[16]

The PINs for TSX and NYSE-listed Canadian stocks are estimated following Easley et al. (1996) and Easley et al. (1997a,b). Further, we adopt Easley et al.'s (2008) log-likelihood function specification for improved numerical stability in computing the the PIN. The bid-ask spreads are adjusted by the mid-quotes and, thus, measure the relative discrepancy between bid and ask quotes free from the exchange rate. Following Eun and Sabherwal (2003), the mid-points of U.S.-Canada exchange rate bid and ask quotes are updated every minute. The bid and ask quotes of the NYSE-listed Canadian stocks are matched with their previous minutes' exchange rate quote mid-points.

[Insert Table 3 about here.]

Table 3 abstracts the exchange-wise estimates and cross-exchange difference tests (Wilcoxon, 1945) of the PIN, relative quoted spread, daily trading volume, and daily trading

---

[14] We owe this comment to Daniel Weaver. See Eun and Sabherwal (2003) for a detailed institutional comparison between the TSX and the NYSE.

[15] See Schultz and Shive (2008) for trade misclassification of the TAQ on the NYSE which becomes severe after 2000.

[16] We owe this comment to Lawrence Kryzanowski. See Fuller, Van Ness, and Van Ness (2008) for difficulties in estimating the PIN for Nasdaq trades.

dollar volume of our sample cross-listed pairs. First, on average, the PIN on the TSX (0.242) exceeds that on the NYSE (0.214). Second, the relative quoted spread on the TSX (0.015) is narrower than that on the NYSE (0.022). Third, for a Canadian cross-listed pair, on average, it appears that the intensity of informed trades tends to be heavier (a higher PIN) with a lower spread (competitive market making) on the TSX.[17] The TSX dominates the NYSE in trading volume in terms of both quantity and value. The statistical significances of comparisons are well explained by all near-zero *p*-values of difference tests.[18]

5.1.2. Estimation of the threshold and convergence speed parameter

In our threshold ECM, the threshold (γ), on average, measures the sum of transaction costs and risk premiums of a cross-border arbitrage. The first column (Threshold) in Panel A of Table 2 reports the summary statistics of estimated thresholds, which range from 0.01 to 0.81, with a mean of 0.193: On average, when the cross-listing dollar premium/discount records more than ±19.3 cents, respectively, arbitrageurs begin to take positions on both sides and drive the deviation back into the "no-arbitrage" band. The percentage of data points falling into the outer regime ranges from 10.24% to 88.69%, with a mean of 17.0% (second column in Panel A of Table 2).

[Insert Table 4 about here.]

According to our smooth transition ECM, the convergence speed parameter (δ), on average, measures the reciprocal speed of converenge to parity of a cross-listed pair (Section

[17] In the cross-section, the PIN is positively correlated with bid-ask spread according to Easley et al. (2002). However, the negative relationship between the two estimates shown in Table 3 is due to averaging and aggregation.

[18] Based on, unreported, ten-minute frequency relative premiums of the cross-listed pairs traded throughout the sample period, the arithmetic mean, the median, and the standard deviation are 0.00306, 0.00004, and 0.03031, respectively. The average relative premium of 30.6 basis points with a 3.03% volatility is a statistically insignificant deviation from parity. This suggests the extent to which Toronto and New York are integrated.

2).[19] In the first column (Delta) of Panel A in Table 4 reports the summary statistics of the estimates of convergence speed parameters where the mean is 0.669 and the median is 0.654. To see how the convergence speed is affected by the price deviation, the second and third columns report the results for δ when there is a premium ($\text{Delta}_{\text{Prem}}$) or discount ($\text{Delta}_{\text{Disc}}$) on the NYSE-listing. Panel B reports the trend of δ and there are downward trends in both mean and median of $\delta,$ which suggest that NYSE and TSX have integrated over time. We apply the Wilcoxon signed rank test to test the null hypothesis in Panel C: $H_0 : \delta_{\kappa<0} \geq \delta_{\kappa>0}$. The *p*-value is smaller than 0.01, thus, we can reject the null hypothesis: The convergence between the two market prices accelerates when there is a relative discount on the NYSE.[20] A possible explanation is that arbitrageurs like to establish short positions on the TSX since stocks are likely to be more liquid in the home market (Table 3).[21]

5.1.3. Estimation of the information share

Given a cross-listed pair, the information share of an exchange measures its contribution to price discovery. We estimate the information shares using the three ECMs described in Section 3. The first column of Panel A in Table 5 reports the estimated information share of NYSE from the standard, linear ECM. With a shorter sample period (February—July, 1998) Eun and Sabherwal's (2003) information share estimates of the NYSE range from 0.2% to 98.2%, with a mean of 38.1% in the cross-section of sample firms. They conclude that price discovery for most cross-listed pairs occurs on the TSX, but there is significant feedback from the NYSE. Our results, based on a longer sample period, are consistent with theirs: The estimated linear information shares of the NYSE (IS) range from 3% to 94.5%, with a mean of 42.99%. There is no discernible trend over the sample period as the

---

[19] The smooth-transition model is estimated nonparametrically without any assumption on the functional form.

[20] Note that convergence is quicker the smaller δ is.

[21] We carefully view this interpretation as valid since the proportion of discounted in cross-listings ($p_2 < p_1$) is 50.6% using 10-minute interval observations through the sample period.

yearly average estimates of the linear information share of NYSE in 1998, 1999, and 2000 are 39.3%, 48.4%, and 41%, respectively (first column in Panel B).

[Insert Table 5 about here.]

However, the linear ECM ignores nonlinearity in the course of parity-convergence, as shown in Subsection 4.2, and this suggest that the estimates from the linear ECM may be biased. Accordingly, we estimate the information shares via both threshold and smooth transition ECMs.

The second and third columns in Panel A of Table 5 report the results for the bivariate threshold ECM presented in Subsection 3.2. The information share estimates of the NYSE in the inner regime (IS_{In}) range from 1.7% to 91%, with an average of 36.2%, in the outer regime (IS_{Out}) range from 2% to 98%, with an average of 43.5%.[22] Thus, overall, the NYSE makes a larger contribution to price discovery in the outer regime: Arbitrageurs tend to engage in the market when price deviations are sizable, and their arbitrage activities can transfer information from the home market to the NYSE (Fremault, 1991).

The last three columns in Panel A of Table 5 report the results from the smooth transition ECM. There are three information shares: IS_{ST}, IS_{Prem}, and IS_{Disc}, which are the information shares defined on the whole sample, sample with premiums and discounts on the NYSE cross-listings, respectively. The means for these three information shares are 37.4%, 38.6%, and 37.9%, respectively. In Panel C, we further apply the Wilcoxon signed rank test to examine the following null hypothesis $H_0: IS_{\kappa<0} \leq IS_{\kappa>0}$. The *p*-value of the test is 0.036, which significantly rejects the null hypothesis. In other words, when the cross-listing trades with a discount on the NYSE, the information share of the NYSE is larger than that of the TSX, which suggests that informed traders may choose to trade on the NYSE when the cross-listing is underpriced than the original listing.

[22] Be reminded that the range of the proportion of outer regime is from 10.2% to 88.7% with a mean of 17%.

5.2. Panel dataset construction

We construct a panel data for regression analyses of the estimates of information shares and thresholds with columns of various indices, dependent variables, explanatory variables, and control variables. *Symbol* is the NYSE ticker of a TSX-NYSE cross-listed pair. *Year* is the year index of an estimated value.

- *Dependent variables. IsOut* and *IsIn* are the outer-regime and inner-regime information shares of the NYSE per threshold ECM. *IsLin* is the information share estimates of the NYSE per Harris et al. (1995, 2002). *Threshold* is the U.S. dollar-denominated threshold estimates. *IsSt*, *IsPrem*, and *IsDisc* are the information shares of the NYSE per smooth trasition ECM using the whole sample, and given premiums and discounts on cross-listings, respectively, per smooth transition ECM.
- *Key explanatory variables.* The PIN captures the informativeness of a listing. We use the estimates of the PINs of both listings of a cross-listed pair to proxy for their relative and average degree of efficient information. Since the informed traders are believed to foster price discovery, the information share of an exchange is expected to be larger relative to its cross-border counterpart the higher the PIN of the listing therein compared to that of its cross-listing. *PINRatio* is the ratio of the PIN of the NYSE over that of the TSX. *PINAvg* is the average PIN of the NYSE and the TSX. *PINDiff* is the difference of the PIN of the NYSE over that of the TSX. We also estimate the relative quoted spread measures on both exchanges to proxy for their respective degrees of market friction. The threshold (effective required return of cross-border arbitrage) of a cross-listed pair is expected to be positively associated with a bid-ask spread measure. *SpreadRatio* is the ratio of the relative quoted bid-ask spread of the NYSE over that of the TSX. *SpreadAvg* is the average relative quoted bid-ask spread of NYSE and the TSX. *SpreadDiff* is the difference of the quoted bid-ask spread of the NYSE over that of the TSX.

- *Control variables. USVol* is the average daily trading volume of the NYSE out of both of the NYSE and the TSX following Eun and Sabherwal (2003). *VolAvg* is the average of the log-transformations of average daily trading volume measures of the NYSE and the TSX. *VolDiff* is the difference of the log-transformation of average daily trading volume of the NYSE over that of the TSX. *USDollarVol* is the average daily dollar trading volume of the NYSE out of both of the NYSE and the TSX. *DollarVolAvg* is the sum of log-transformations of average daily dollar trading volume measures of the NYSE and the TSX. *DollarVolDiff* is the difference of the log-transformation of average daily dollar trading volume of the NYSE over that of the TSX. *Governance* is the Report on Business governance index of Canadian firms published by *Globe and Mail* (McFarland, 2002). *Industry* equals one if the cross-lister is a manufacturing firm, and zero otherwise. We believe the governance risk of a Canadian firm is reflected in the threshold as a risk premium. *Size* is the normalized average market capitalization on the TSX and the NYSE.

5.3. Panel regression analyses

5.3.1. Regressions of the information shares per threshold cointegration

We conduct regression analyses on the constructed panel data to identify the factors that affect the relative extent of the NYSE's contribution to price discovery. The estimated outer-regime information shares are regressed onto the panel of explanatory and control variables and reported in Panel A of Table 6, respectively. It turns out that the contribution of the NYSE increases relatively against that of the TSX as the NYSE-based trades become more informative (PIN). This is cross-border evidence that informed trades contribute to fostering price discovery, in line with Chen and Choi (2012). Either in quantity or value, the higher the liquidity on the NYSE the more it leads in price discovery. This is consistent with Eun and Sabherwal's (2003) findings: they estimate the information share of the NYSE by using Harris et al.'s (1995, 2002) approach. They find that the information share is directly related to the U.S.'s share of total trading (UsVol), the proportion of informative trades on U.S. exchanges and the TSX (confirmed as proxied by the PIN), and the inversely related to the ratio of bid-ask spreads on U.S. exchanges and the TSX, which is not discernable.[23] A better investor-protecting (Governance) and larger (Size) Canadian firm tends to lead price setting on the TSX as seen in Models 1 through 4.

[Insert Table 6 about here.]

We conduct analoguous panel regressions for the inner-regime and linear information shares in Panels B and C of Table 6, respectively. Neither alternative measure of exchange-specific contribution to price dicovery has a higher explanatory power (adjusted R^2) and economically and statistically meaningful implications. From this end, the outer-regime information shares (Panel A) have not only proved heuristically appealing but also economically reasonable and statistically robust.

5.3.2. Regression of the estimated threshold

For each cross-listed pair, the threshold includes transactions costs, which consist of bid-ask price spreads on both exchanges and the foreign exchange rate, fixed costs, and liquidity shorfalls. Implicit risk premiums, including those from information asymmetry and macroeconomic uncertainty, can also affect the determination of the threshold. Accordingly, Table 7 provides the results of panel regressions of the estimated thresholds onto average (Panel A) and difference (Panel B) measures of asymmetric information component (PIN) and the inverse of market depth (spread), controlling for liquidity, either in quantity (UsVol) or value (UsDollarVol), firm-level idiosyncratic characteristics (Industry, Governance, and

[23] Hasbrouck (1995) finds that there is a positive and significant correlation between contribution to price discovery made by the NYSE and its market share by trading volume using the U.S. domestic data. Using the same data, Harris et al. (2002) finds evidence that the information share increases when its bid-ask spreads decline relative to the regional exchange.

Size), and interest rates (yields of 90-day bills and 10-year notes).

[Insert Table 7 about here.]

As expected, our measure of market friction (relative quoted spread) significantly increases the required dollar return of cross-border arbitrage as 4 out of 8 models using average measures (Panel A) and all models using difference measures (Panel B) agree with it. The better the firm is governed at home, the lower the minimum required profit as all models with the Governance control variable show. Manufacturing firms (when Industry equals 1) tend to require larger relative premiums to be exploited. Overall, difference measures turn out to have a greater determination on the threshold level than the average measures do as the adjusted R^2's of Panel B dominate those of Panel A through all specifications. In sum, the effective break-even point (threshold) of cross-border arbitrage appears to be affected by the relative degree of private information, market friction, and liquidity measures, and idiosyncratic firm-level characteristics. These, much economically appealing, empirical results lend support to the findings of Gagnon and Karolyi (2010).

5.3.3. Regressions of the information shares per smooth trasition cointegration

Since parity-convergence of a cross-listed pair can also be gradual rather than abrupt, an alternative measure of contribution to price discovery is the smooth transition ECM-implied information share as we proposed (Subsection 3.3) and estimated (Subsection 5.1.3) earlier: IsSt, IsPrem, and IsDisc are the information shares of NYSE using the whole sample, and given premiums and discounts on cross-listings, respectively. These information shares are dependent variables in the panel regressions onto the same group of explanatory and control variables shown in the regressions of threshold ECM-implied information shares (Table 6) whose respective results are shown in Panels A, B, and C of Table 8. We conduct separate regressions of information shares given premiums (Panel B) versus discounts (Panel C) on cross-listings since the NYSE is shown to be dominant in contribution to price discovery in the former case (Table 5 Panel C).

[Insert Table 8 about here.]

In Panel A, the contribution of the NYSE to the price discovery of a cross-listed pair, assuming gradual convergence to parity, appears to be more influential with a higher relative population of informed traders (PinRatio), market friction (SpreadRatio), and liquidity in both quantity (UsVol) and value (UsDollarVol) on the NYSE vis-à-vis TSX in Models 1 through 6. Compared to the regression results of the outer-regime information share of the NYSE (Table 6 Panel A), the corporate governance (Governance) of the Candian cross-lister does not appear to be effective in determining the venue of price discovery when convergence is gradual (smooth transition cointegration) rather than abrupt (threshold cointegration). Overall, the smooth transition ECM-implied information share appears to be explained by a similar basket of risk factors and controls as in the case of threshold ECM-implied outer-regime information share. The model fitness of both measures of information share is also deemed comparable: The adjusted R^2's of smooth transition ECM regression models (Table 8 Panel A) versus that of threshold ECM (Table 6 Panel A) are in the range of 9.6%—29.7% versus 14.4%—27.7%, respectively. These results are in stark comparison with those of linear ECM-implied information share (Table 6 Panel C) both in terms of statistical and economic significances of risk factors and controls, and model fitness. Thus, not only there exists undeniable nonlinearity (Table 2) in the course of convergence to parity, but the contribution measures of price discovery are better explained when nonlinearity is assumed, either threshold regime switch or smooth transition. Panels B and C report qualitatively similar regression outcomes whether we use the information share given premiums or discounts on cross-listings, respectively.

6. Conclusion

For a pair of the original listing and its cross-listing, the adjustment to parity can be discontinuous: Convergence may be quicker when the relative premium is profitable, or slower otherwise. In other words, the dynamics of cross-listed pairs fall into two regimes: Within and beyond the threshold, e.g. transaction costs and associated risk premiums of arbitrage. This paper extends Harris et al.'s (1995, 2002) ECM to estimate the extent of contribution to price discovery (information share) by considering threshold cointegration per Balke and Fomby (1997). Alternatively, since convergence may occasionally be gradual and nonlinear, we further generalize the threshold framework to a smooth transition version.

According to our threshold and smooth transition ECMs, the information share and threshold are estimated and regressed with following empirical implications: First, parity-convergence accelerates upon discounts on the cross-listings on the NYSE. Second, we find a larger feedback from the NYSE if the price gap exceeds the threshold. Third, informed traders tend to cluster on the NYSE upon discounts on the cross-listings. Fourth, the information share and threshold estimates are affected by the relative degree of private information, market friction and liquidity measures, and firm-level characteristics.

Lastly, as a disclaimer, we do not account for exchange-rate market friction in our threshold ECM framework unlike Grammig et al. (2005). This is because of the stationarity of U.S.-Canada exchange rate (Issa et al., 2006), and that the synchronous trading environment of TSX-NYSE cross-listed pairs allows constructing a cointegration system without considering the exchange-rate bid-ask spread which risks a sufficiently low margin of error (Eun and Sabherwal, 2003). However, introducing a such additional source of randomness to modeling nonlinear dynamics of cross-listed stocks should be of interest for future studies.

Acknowledgements

Choi is grateful for Kwanjeong Educational Foundation grants 03112BUS044 and 0422USD012. Special thanks are due to Warren B. Bailey and G. Andrew Karolyi. We would also like to thank Jinho Byun, Michael Chng, Stefan Frey, Jinho Kim, Yihui Lan, Lars Norden, Alexei Orlov, Jung Soon Shin, Jungwon Suh, Yafeng Qin, Byung Sam Yoo and the participants in INFINITI 2010 (Dublin, Ireland), Asian Finance Association 2010 (Hong Kong), Securities and Financial Markets 2011 (Kaohsiung, Taiwan), and Australian Finance and Banking Conference 2011 (Sydney, Australia) for invaluable discussions and feedback. The paper has previously circulated under the title: "Price Discovery and Threshold Cointegration: Theory and Empirics on Cross-Border Stock Trading." Standard disclaimer rules apply and all errors are of our own.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716-723.

Bacidore J., Sofianos G., 2002. Liquidity provision and specialist trading in NYSE-listed non-U.S. stocks. Journal of Financial Economics 63, 133-158.

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. International Economic Review 38, 627-645.

Chan, K., Menkveld, A.J., Yang, Z., 2008. Information asymmetry and asset prices: Evidence from the China foreign share discount. The Journal of Finance 63, 159-196.

Chen, H., Choi, P.M.S., 2012. Does information vault Niagara Falls? Cross-listed trading in New York and Toronto. Journal of Empirical Finance, Forthcoming.

Chen,Y.L., Gau, Y.F., 2010. News announcements and price discovery in foreign exchange spot and futures markets. Journal of Banking and Finance 34, 1628-1636.

Choi, I., 2001. Unit root tests for panel data. Journal of International Money and Finance 20, 249-272.

Chung, H., Ho, T.-W., Wei, L.-J., 2005. The dynamic relationship between the prices of ADRs and their underlying stocks: Evidence from the threshold vector error correction model. Applied Economics 37, 2387-2394.

De Jong, F., 2002. Measures of contributions to price discovery: A comparison. Journal of Financial Markets 5, 323-327.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427--431.

Doidge, C., Karolyi, G.A., Stulz, R.M., 2009. Has New York become less competitive in global markets? Evaluating foreign listing choices over time. Journal of Financial Economics 91, 253-287.

Duarte, J., Young, L., 2009. Why is PIN priced? Journal of Financial Economics 91, 119-138.

Easley, D., Engle, R.F., O'Hara, M., Wu, L., 2008. Time-varying arrival rates of informed and uninformed trades. Journal of Financial Ecnometrics 6, 171-207.

Easley, D., Hvidkjaer, S., O'Hara, M., 2002. Is information risk a determinant of asset returns? The Journal of Finance 57, 2185-2221.

Easley, D., Kiefer, N.M., O'Hara, M., 1997. One day in the life of a very common stock. Review of Financial Studies 10, 805-835.

Easley, D., Kiefer, N.M., O'Hara, M., 1997. The information content of the trading process. Journal of Empirical Finance 4, 159-186.

Easley, D., Kiefer, N.M., O'Hara, M., Paperman, J.B., 1996. Liquidity, information, and infrequently traded stocks. The Journal of Finance 51, 1405-1436.

Easley, D., O'Hara, M., Saar, G., 2001. How stock splits affect trading: A microstructure approach. Journal of Financial and Quantitative Analysis 36, 25-51.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: Representation, estimation and testing. Econometrica 35, 251-276.

Engle, R.F., Yoo, B.S., 1987. Forecasting and testing in co-integrated systems. Journal of Econometrics 35, 143-159.

Eun, C., Sabherwal, S., 2003. Cross-border listings and price discovery: Evidence from U.S.-listed Canadian stocks. The Journal of Finance 58, 549-574.

Figuerola-Ferretti, I., Gonzalo, J., 2009. Modeling and measuring price discovery in commodity markets. Journal of Econometrics, forthcoming.

Fremault, A., 1991. Stock index futures and index arbitrage in a rational expectation model. Journal of Business 64, 523-547.

Fricke, C., Menkhoff, L., 2011. Does the "Bund" dominate price discovery in Euro bond futures? Examing information shares. Journal of Banking and Finance 35, 1057-1072.

Fuller, K., Van Ness, B., Van Ness, R. Is information risk priced for NASDAQ-listed securities? Working paper, University of Mississippi.

Gagnon, L., Karolyi, G.A., 2010. Multi-market trading and arbitrage. Journal of Financial Economics 97, 53-80.

Garbade, K.D., Silber, W.L., 1983. Price movements and price discovery in futures and cash markets. The Review of Economics and Statistics 65, 289-297.

Gonzalo, J., Granger, C.W.J., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics 13, 1-9.

Gonzalo, J., Martinez, O., 2006. Large shocks vs small shocks (or does size matter? May be so). Journal of Econometrics 135, 311-347.

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.

Grossman, S., Stiglitz, J, 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393-408.

Hansen, B.E., Seo, B., 2002. Testing for two-regime threshold cointegration in vector error correction models. Journal of Econometrics 110, 293-318.

Harris, F.H.deB., McInish, T.H., Shoesmith, G., Wood, R.A., 1995. Cointegration, error correction, and price discovery on informationally linked security markets. Journal of Financial and Quantitative Analysis 30, 563-579.

Harris, F.H.deB., McInish, T.H., Wood, R.A., 1997. Common long-memory components of intraday stock prices: A measure of price discovery. Working paper, Wake Forest University.

Harris, F.H.deB., 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, 1175-1201.

Hasbrouck, J., 2002. Stalking the efficient price in market microstructure specifications: An overview. Journal of Financial Markets 5, 329-339.

Issa, R., Lafrance, R., Murray, J., 2006. The turning black tide: Energy prices and the Canadian dollar. Working paper, Bank of Canada.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551-1580.

Kim, M., Szakmary, A.C., Mathur, I., 2000. Price transmission dynamics between ADRs and their underlying foreign securities. Journal of Banking and Finance 24, 1359-1382.

Koumkwa, S., Susmel, R., 2005. Arbitrage and convergence: evidence from Mexican ADRs. Working paper, University of Houston.

Krugman, P.R., 1991. Target zones and exchange rate dynamics. Quarterly Journal of Economics 106, 669-682.

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

Lee, C.M.C., Ready, M.J., 1991. Inferring trade direction from intraday data. The Journal of Finance 46, 733-746.

Liu, Q., An, Y., 2011. Information transmission in informationally linked markets: Evidence from US and Chinese Commodity futures markets. Journal of International Money and Finance 30, 778-795.

McFarland, J., 2002. Report of Business. Globe and Mail, 7 October, B6.

Neal, R. 1996. Direct tests of index arbitrage models. The Journal of Financial and Quantitative Analysis 31, 541-562.

Rabinovitch, R., Silva, A.C., Susmel, R., 2003. Impact of capital controls and transaction costs on the return distribution of dually traded securities: Evidence from Chile and Argentina. Working paper, University of Houston.

Schreiber, P.S., Schwartz, R.A., 1986. Price discovery in securites markets. Journal of Portfolio Management 12, 43-48.

Schultz, P., Shive, S., 2008. Mispricing of dual class shares: Profit opportunities, arbitrage, and trading. Working paper, University of Notre Dame.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461-464.

Solnik, B., Boucrelle, C., Le Fur, Y., 1996. International market correlation and volatility. Financial Analysts Journal 52, 17-34.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics 1, 80-83.

Appendix A. Estimation and testing of parameters

For convenience, the firm indicator (i) is omitted in the following discussion. The threshold ECM aforementioned in Subsection 3.2 can be represented as follows:

$$\Delta x_t = A_1^T X_{t-1} d_{1t}(\gamma) + A_2^T X_{t-1} d_{2t}(\gamma) + u_t, \tag{A.51}$$

where $\Delta x_t = (p_{1t}, p_{2t})$, $X_{t-1} = [1, \kappa_{t-1}, \Delta x_{t-1}, \Delta x_{t-2}, ..\Delta x_{t-m}]^T$, $d_{1t}(\gamma) = \mathbf{1}(|\kappa_{it-1}| \leq \gamma_i)$ and $d_{1t}(\gamma) = \mathbf{1}(|\kappa_{t-1}| > \gamma)$; A_1^T and A_2^T contain the parameters to be estimated; and γ is the threshold parameter to be estimated.

The threshold VECM can be estimated using the MLE method proposed by Hansen and Seo (2002). Assuming that the error term (u_t) is *i.i.d.* Gaussian, the likelihood function is

$$\mathcal{L}_n(A_1, A_2, \Sigma, \gamma) = -\frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{n} u_t(A_1, A_2, \gamma)^T \Sigma^{-1} u_t(A_1, A_2, \gamma), \tag{A.52}$$

where $u_t(A_1, A_2, \gamma) = \Delta x_t - A_1^T X_{t-1} d_{1t}(\gamma) - A_2^T X_{t-1} d_{2t}(\gamma)$. The covariance matrix (Σ) is an identity matrix due to the *i.i.d.* Gaussian assumption of the error term. For a fixed c, A_1 and A_2 can estimated by an OLS regression, thus

$$\hat{A}_1(\gamma) = \left(\sum_{t=1}^{n} X_{t-1} X_{t-1}^T d_{1t}(\gamma)\right)^{-1} \sum_{t=1}^{n} X_{t-1} \Delta x_t^T d_{1t}(\gamma), \tag{A.53}$$

$$\hat{A}_2(\gamma) = \left(\sum_{t=1}^{n} X_{t-1} X_{t-1}^T d_{2t}(\gamma)\right)^{-1} \sum_{t=1}^{n} X_{t-1} \Delta x_t^T d_{2t}(\gamma), \tag{A.54}$$

and then $\hat{u}_t(\gamma) = \Delta x_t - \hat{A}_1^T X_{t-1} d_{1t}(\gamma) - \hat{A}_2^T X_{t-1} d_{2t}(\gamma)$. By plugging $\hat{u}_t(\gamma)$, the likelihood function $\left(\mathcal{L}_n(A_1, A_2, \Sigma, \gamma)\right)$ becomes a univariate function of γ:

$$\mathcal{L}_n(\gamma) = \frac{-n}{2}\ln\left(\frac{1}{n}\sum_{t=1}^{n} \hat{u}_t(\gamma)\hat{u}_t(\gamma)^T\right) - \frac{n(m+2)}{2}. \tag{A.55}$$

Following Hansen (2000), the *grid search* method can be used to estimate γ within an preset interval $[\underline{\gamma}, \overline{\gamma}]$. The mle estimators for A_1 and A_2 can be obtained by inserting $\hat{\gamma}$. To further confirm the threshold effect, we need to test the following null hypothesis:

$$H_0: A_1 = A_2 \text{ for any } \gamma \in [\underline{\gamma}, \overline{\gamma}] \tag{A.56}$$

against

$$H_1: A_1 \neq A_2 \text{ for some } \gamma \in [\underline{\gamma}, \overline{\gamma}]. \tag{A.57}$$

We use the supremum-Lagrangian multiplier (supLM) test (Hansen and Seo, 2002) to test the above hypotheses. The LM statistic is

$$\mathcal{LM}(\gamma) = \left(\hat{A}_1(\gamma) - \hat{A}_2(\gamma)\right)^T \left(\hat{V}_1(\gamma) + \hat{V}_2(\gamma)\right)^{-1} \left(\hat{A}_1(\gamma) - \hat{A}_2(\gamma)\right), \tag{A.58}$$

where $\hat{V}_1(\gamma) = M_j(\gamma)^{-1}\Omega_j(\gamma)M_j(\gamma)^{-1}, M_j(\gamma) = I_{m+2} \otimes \Pi_j(\gamma)^T\Pi_j(\gamma);$ and$\Omega_j(\gamma) = \Gamma_j(\gamma)^T\Gamma_j(\gamma)$, and $\Pi_j(\gamma), \Gamma_j(\gamma)$ are matrices of the stacked rows of $X_{t-1}d_{jt}(\gamma)$ and $\hat{u}_t(\gamma) \otimes X_{t-1}d_{jt}(\gamma)$, respectively. Define

$$\sup\mathcal{LM} = \sup_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\mathcal{LM}(\gamma). \tag{A.59}$$

A bootstrap method is used to generate the critical value since the asymptotic distribution is non-standard.

Table 1

Estimated cointegrating vector.

The prices of the sample TSX-NYSE Cross-listed pairs are tested for cointegration per Johansen (1991). Since we have only two price series in each regression equation in the cointegrated system, there is at most one cointegrating vector. We estimate the normalized cointegrating vector $(1,-b)^T$ by each firm-year. Our results show that most of the estimated cointegrating vectors are $(1,-1)^T$, which is of the expected values according to the law of one price. The t-statistics for the null hypothesis attests that the cointegrating vector equals $(1,-1)^T$. The observations are in firm-years.

	b	*t* -statistic
5 %-ile	0.9	-5.25
25 %-ile	0.995	-1.29
Median	0.999	0.25
75 %-ile	1.002	0.99
95 %-ile	1.011	2.94

Table 2
Nonlinearity tests.

We estimate the threshold (required cross-border arbitrage return) per our threshhold ECM framework following Balke and Fomby (1997) and extended from Harris et al. (1995, 2002). supLM is the threshold effect (supremum Lagrangian multiplier) test statistic estimated per Hansen and Seo (2002). In order to examine whether the threshold effect happens on the coefficients of error correction term or short dynamic term in the thresholdhold ECM, $Wald_{ECM1}$ and $Wald_{ECM2}$ are the Wald test statistics for the null hypotheses of "no threshold effect" on the error correction terms for TSX and NYSE lisstings, respectively. $Wald_{DC1}$ and $Wald_{DC2}$ are the Wald test statistics for the the null hypotheses of "no threshold effect" on the short-term dynamics terms for TSX and NYSE lisstings, respectively.

Panel A: Threshold estimates and supLM test statistics.

	Threshold	Outer regime, %-age	supLM	95%-ile critical value.	*p* -value
Mean	0.193	0.170	73.516	24.783	0.146
St. Dev.	0.159	0.162	110.814	3.562	0.249
1%-ile	0.010	0.102	10.290	17.861	0.000
10%-ile	0.059	0.103	14.142	21.113	0.000
25%-ile	0.100	0.103	19.664	21.920	0.000
50%-ile	0.157	0.108	31.054	23.101	0.008
75%-ile	0.242	0.143	56.139	28.234	0.142
90%-ile	0.319	0.279	242.699	29.078	0.609
99%-ile	0.808	0.887	509.250	30.407	0.818

Panel B: Wald statistics.

	$Wald_{ECM1}$	*p* -value	$Wald_{ECM2}$	*p* -value	$Wald_{DC1}$	*p* -value	$Wald_{DC2}$	*p* -value
Mean	10.118	0.265	32.723	0.260	13.330	0.265	9.349	0.246
St. Dev.	24.471	0.312	87.764	0.324	70.767	0.283	14.630	0.281
1%-ile	0.000	0.000	0.000	0.000	0.158	0.000	0.183	0.000
10%-ile	0.099	0.000	0.054	0.000	0.796	0.002	1.162	0.001
25%-ile	0.440	0.005	0.336	0.000	2.012	0.033	2.222	0.028
50%-ile	2.637	0.104	3.193	0.074	5.676	0.167	5.554	0.108
75%-ile	8.068	0.507	16.106	0.563	9.679	0.428	9.636	0.400
90%-ile	15.839	0.755	75.568	0.818	13.810	0.723	14.601	0.759
99%-ile	132.409	0.984	487.706	0.986	28.317	0.981	62.062	0.920

Table 3

Representative statistics of key independent variables.

PIN is the probability of information-based trading per Easley et al. (1996). Spread is the relative quoted spread: Bid-ask spread divided by the quoted mid point. Volume is the total daily trading volume in quantity. DollarVol is the total daily trading volume in value.

Variable	Exchange	Mean	Median	St. Dev.	Firm-years	Hypothesis	p-value
PIN	TSX	0.242	0.213	0.107	104	H_0 : $PIN_{TSX} \leq PIN_{NYSE}$	0.001
	NYSE	0.214	0.202	0.060	104	H_1 : $PIN_{TSX} > PIN_{NYSE}$	
Spread	TSX	0.015	0.007	0.025	104	H_0 : $Spread_{TSX} \geq Spread_{NYSE}$	0.000
	NYSE	0.022	0.015	0.022	104	H_1 : $Spread_{TSX} < Spread_{NYSE}$	
Volume (× 1,000)	TSX	576.458	272.687	937.816	104	H0 : $Volume_{TSX} \leq Volume_{NYSE}$	0.000
	NYSE	276.955	59.472	847.500	104	H1 : $Volume_{TSX} > Volume_{NYSE}$	
Dollar Vol. (× $\$10^6$)	TSX	17.236	5.543	43.032	104	H_0 : $DollarVol_{TSX} \leq DollarVol_{NYSE}$	0.001
	NYSE	11.987	1.153	53.373	104	H_1 : $DollarVol_{TSX} > DollarVol_{NYSE}$	

Table 4
Delta estimates of NYSE.

According to the smooth transition ECM, Delta of a cross-listed pair is the convergence speed parameter which, on average, measures the reciprocal speed of converenge to parity. DeltaPrem and DeltaDisc are the convergence speed parameters given premiums and discounts on NYSE-cross-listings, respectively.

Panel A: Representative statistics.

	Delta	$Delta_{Prem}$	$Delta_{Disc}$
Mean	0.669	0.688	0.652
St. Dev.	0.105	0.133	0.123
1%	0.495	0.446	0.434
10%	0.550	0.537	0.532
25%	0.585	0.603	0.580
50%	0.654	0.661	0.641
75%	0.734	0.760	0.722
90%	0.827	0.883	0.814
99%	0.897	1.000	0.894

Panel B: Annual mean estimates.

	Delta		$Delta_{Prem}$		$Delta_{Disc}$	
	Mean	Median	Mean	Median	Mean	Median
1998	0.709	0.709	0.729	0.724	0.701	0.713
1999	0.653(0.08)	0.653	0.668(0.12)	0.644	0.641(0.09)	0.616
2000	0.65(0.67)	0.650	0.674(0.17)	0.616	0.619(0.61)	0.620

Panel C: Wilcoxon signed rank test.

Hypothesis	Wilcoxon Stat.	p-value
$H_0 : Delta_{Disc} \geq Delta_{Prem}$ $H_1 : Delta_{Disc} < Delta_{Prem}$	1,312	6.446×10^{-05}

Note: Panel B, the numbers in the brackets are
p-value to the t test for the significance of difference between the current year and previous year

Table 5
Information shares of NYSE.

The information share of the NYSE is a relative measure of contribution made by the NYSE to price discovery of TSX-NYSE cross-listed pairs. IS is the linear information share following Harris et al.'s (1995, 2002) standard ECM. Per threshold ECM, IS_{In} and IS_{Out} are the inner and outer-regime information shares of the NYSE. Per smooth transition ECM, IS_{ST}, IS_{Prem}, and IS_{Disc} which are the information shares of NYSE using the whole sample, and given premiums and discounts on cross-listings, respectively.

Panel A: Summary statistics.

	Linear ECM	Threshold ECM		Smooth Transition ECM		
	IS	IS_{In}	IS_{Out}	IS_{ST}	IS_{Prem}	IS_{Disc}
Mean	0.430	0.362	0.435	0.374	0.386	0.379
St. Dev.	0.258	0.239	0.259	0.254	0.253	0.264
1%-ile	0.030	0.017	0.020	0.001	0.012	0.003
10%-ile	0.087	0.073	0.106	0.059	0.067	0.068
25%-ile	0.215	0.138	0.215	0.173	0.177	0.161
50%-ile	0.416	0.358	0.418	0.352	0.369	0.360
75%-ile	0.601	0.543	0.626	0.543	0.554	0.536
90%-ile	0.816	0.669	0.804	0.707	0.739	0.797
99%-ile	0.948	0.910	0.980	0.946	0.934	0.981

Panel B: Annual mean estimates.

	Linear ECM	Threshold ECM		Smooth Transition ECM		
	IS	IS_{In}	IS_{Out}	IS_{ST}	IS_{Prem}	IS_{Disc}
1998	0.393	0.367	0.386	0.386	0.413	0.381
1999	0.484	0.368	0.514	0.382	0.378	0.401
2000	0.410	0.352	0.442	0.357	0.350	0.374

Panel C: Wilcoxon signed rank test of smooth transition information share.

Hypothesis	Wilcoxon Stat.	p-value
$H_0 : IS_{Disc} \leq IS_{Prem}$ $H_1 : IS_{Disc} > IS_{Prem}$	2,877	0.036

Table 6

Panel regression results of threshold and linear ECM-implied information shares.

The dependent variables of Panels A, B, and C are IsOut which is the outer-regime information share of the NYSE which is a relative measure of contribution made by the NYSE to price discovery of TSX-NYSE cross-listed pairs; IsIn which is the inner-regime information share of the NYSE; and IsLin which is the linear information share (Harris et al., 1995, 2002). Explanatory variables are: PinRatio is the ratio of the PIN of the NYSE over that of the TSX. SpreadRatio is the ratio of the relative quoted bid-ask spread of the NYSE over that of the TSX. UsVol is the average daily trading volume of the NYSE out of both of the NYSE and the TSX following Eun and Sabherwal (2003). UsDollarVol is the average daily dollar trading volume of the NYSE out of both of the NYSE and the TSX. Control variables. Governance is the Report on Business governance index of Canadian firms published by Globe and Mail (McFarland, 2002). Industry equals one if the cross-lister is a manufacturing firm, and zero otherwise. Size is the normalized average market capitalization on the TSX and the NYSE. The t-statistics of coefficient estimates are suppresed for lack of space. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively. The observations are in firm-years. All specificiations are controlled for fixed and year effects.

Panel A: Outer-regime information shares.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	0.651 ***	0.702 ***	0.632 ***	0.683 ***	0.262 ***	0.307 ***	0.206 ***	0.242 ***
PinRatio	0.127 **	0.122 **	0.133 **	0.127 **	0.151 ***	0.136 **	0.179 ***	0.168 ***
SpreadRatio	0.001	0.002	0.002	0.002	0.000	-0.001	-0.002	-0.002
UsVol	0.386 ***		0.358 ***		0.414 ***		0.454 ***	
UsDollarVol		0.300 ***		0.277 ***		0.282 ***		0.336 ***
Industry	-0.054	-0.050						
Governance	-0.005 ***	-0.005 ***	-0.005 ***	-0.005 ***				
Size	-0.390 **	-0.403 **	-0.353 **	-0.368 **	-0.443 **	-0.473 **		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	0.277	0.252	0.273	0.249	0.207	0.154	0.203	0.144

Panel B: Inner-regime information shares.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.027	-0.026	-0.027	-0.026	-0.020	-0.018	-0.022 *	-0.021
PinRatio	-0.015	-0.016	-0.015	-0.016	-0.007	-0.008	0.008	0.007
SpreadRatio	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	-0.001
UsVol	0.225		0.225		0.200		0.234 *	
UsDollarVol		0.222		0.222		0.213		0.247 *
Industry	0.000	-0.001						
Governance	0.000	0.000	0.000	0.000				
Size	-0.030	-0.036	-0.029	-0.035	-0.033	-0.036		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	0.015	0.014	0.025	0.023	0.018	0.018	0.044	0.044

Panel C: Linear information shares.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	0.015	0.014	0.013	0.011	0.014	0.014	0.019	0.019
PinRatio	0.049	0.055	0.049	0.055	0.052	0.057	0.065	0.071
SpreadRatio	0.002	0.001	0.002	0.001	0.002	0.001	0.001	0.001
UsVol	-0.153		-0.151		-0.128		-0.034	
UsDollarVol		-0.350		-0.348		-0.330		-0.189
Industry	-0.004	-0.005						
Governance	0.000	0.000	0.000	0.000				
Size	-0.030	-0.045	-0.025	-0.040	-0.019	-0.034		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	-0.014	0.006	-0.005	0.014	0.010	0.029	0.017	0.025

Table 7
Panel regression results of threshold values.

The dependent variable is Threshold which is the U.S.$-denominated threshold estimate. Explanatory variables: PinDiff is the difference of the PIN of the NYSE over that of the TSX. PinAvg is the average PIN of the NYSE and the TSX. SpreadDiff is the difference of the quoted bid-ask spread of the NYSE over that of the TSX. SpreadAvg is the average relative quoted bid-ask spread of NYSE and the TSX. Control variables: VolAvg is the average of the log-transformations of average daily trading volume measures of the NYSE and the TSX. VolDiff is the difference of the log-transformation of average daily trading volume of the NYSE over that of the TSX. DollarVolAvg is the sum of log-transformations of average daily dollar trading volume measures of the NYSE and the TSX. DollarVolDiff is the difference of the log-transformation of average daily dollar trading volume of the NYSE over that of the TSX. Governance is the Report on Business governance index of Canadian firms published by Globe and Mail (McFarland, 2002). Industry equals one if the cross-lister is a manufacturing firm, and zero otherwise. Size is the normalized average market capitalization on the TSX and the NYSE. NoteAvg and NoteDiff are the average and difference of US and Canada's 10-year Treasury Note yields, respectively. BillAvg and BillDiff are the average and difference of US and Canada's 90-day Treasury bill discounts, respectively. VolatAvg and VolatDiff are the average and difference of US and Canada's market index return volatility, respectively. GdpAvg and GdpDiff are the average and difference of US and Canada's GDP growth rates, respectively. The t-statistics of coefficient estimates are suppresed for lack of space. The observations are in firm-years. All model specifications are controlled for fixed and year effects.

Panel A: Regressions onto average measures.

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8	
Intercept	1.275		2.488	*	1.085		2.591	*	0.373		1.852		0.625		1.742	*
PinAvg	-1.419		-2.152		-1.087		-2.082		-0.053		-0.945		-0.410		-1.131	
SpreadAvg	15.217	***	11.387	*	15.419	***	11.923	*	3.959		0.782		2.789		-0.214	
VolAvg	0.003				0.032				0.024				0.008			
DollarVolAvg			-0.066				-0.060				-0.067				-0.056	
Industry	0.366	***	0.370	***												
Governance	-0.010	***	-0.011	***	-0.010	**	-0.010	**								
Size	0.458		0.789		0.013		0.411		-0.290		0.126					
No. of Obs.	104		104		104		104		104		104		104		104	
Adjusted R^2	0.118		0.126		0.048		0.052		-0.034		-0.027		-0.029		-0.021	

Panel B: Regressions onto difference measures.

	Model 9		Model 10		Model 11		Model 12		Model 13		Model 14		Model 15		Model 16	
Intercept	1.031	***	1.007	***	1.278	***	1.268	***	0.567	***	0.574	***	0.589	***	0.590	***
PinDiff	-1.553	*	-1.427	*	-1.731	*	-1.462		-1.206		-1.212		-1.067		-1.048	
SpreadDiff	10.461	***	9.386	***	10.091	***	10.050	***	9.299	***	9.115	***	7.959	***	8.064	***
VolDiff	-0.093	**			-0.051				-0.013				0.002			
DollarVolDiff			-0.065	*			-0.019				-0.011				0.004	
Industry	0.495	***	0.491	***												
Governance	-0.013	***	-0.011	***	-0.011	***	-0.010	**								
Size	0.194		0.192		-0.170		-0.132		-0.315		-0.318					
No. of Obs.	104		104		104		104		104		104		104		104	
Adjusted R^2	0.208		0.193		0.086		0.076		0.031		0.031		0.035		0.036	

Table 8

Panel regression results of smooth transition ECM information shares.

The dependent variables of Panels A, B, and C are, per smooth transition ECM, IsSt, IsPrem, and IsDisc which are the information shares of NYSE using the whole sample, and given premiums and discounts on cross-listings, respectively. Explanatory variables are: PinRatio is the ratio of the PIN of the NYSE over that of the TSX. SpreadRatio is the ratio of the relative quoted bid-ask spread of the NYSE over that of the TSX. UsVol is the average daily trading volume of the NYSE out of both of the NYSE and the TSX following Eun and Sabherwal (2003). UsDollarVol is the average daily dollar trading volume of the NYSE out of both of the NYSE and the TSX. Control variables. Governance is the Report on Business governance index of Canadian firms published by Globe and Mail (McFarland, 2002). Industry equals one if the cross-lister is a manufacturing firm, and zero otherwise. Size is the normalized average market capitalization on the TSX and the NYSE. The t-statistics of coefficient estimates are suppresed for lack of space. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively. The observations are in firm-years. All specificiations are controlled for fixed and year effects.

Panel A. Information shares with whole sample.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.121	-0.122	-0.060	-0.060	-0.066	-0.066	-0.025	-0.025
PinRatio	0.120 **	0.120 **	0.119 ***	0.120 ***	0.124 ***	0.124 ***	0.064	0.064
SpreadRatio	0.096 ***	0.097 ***	0.098 ***	0.099 ***	0.099 ***	0.100 ***	0.055 *	0.056 *
UsVol	0.726 ***		0.727 ***		0.735 ***		0.431 *	
UsDollarVol		0.726 ***		0.727 ***		0.735 ***		0.448 *
Industry	0.016	0.016						
Governance	0.000	0.000	0.000	0.000				
Size	0.001	0.001	0.000	0.000	0.001	0.001		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	0.285	0.285	0.293	0.293	0.297	0.296	0.096	0.098

Panel B. Information shares given premiums on cross-listings.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.049	-0.050	-0.023	-0.023	-0.025	-0.025	-0.003	-0.003
PinRatio	0.081 *	0.081 *	0.081 *	0.081 *	0.084 *	0.084 *	0.051	0.050
SpreadRatio	0.117 ***	0.118 ***	0.118 ***	0.119 ***	0.118 ***	0.118 ***	0.067 **	0.067 **
UsVol	0.431		0.432		0.443 *		0.120	
UsDollarVol		0.433		0.434		0.446 *		0.139
Industry	0.007	0.007						
Governance	0.000	0.000	0.000	0.000				
Size	0.001	0.001	0.000	0.000	0.000	0.000		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	0.249	0.249	0.259	0.260	0.270	0.270	0.056	0.057

Panel C. Information shares given discounts on cross-listings.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.165	-0.166	-0.082	-0.083	-0.091	-0.091	-0.043 *	-0.043 *
PinRatio	0.147 ***	0.147 ***	0.146 ***	0.147 ***	0.153 ***	0.154 ***	0.098 **	0.099 **
SpreadRatio	0.096 ***	0.096 ***	0.098 ***	0.099 ***	0.100 ***	0.100 ***	0.071 **	0.072 **
UsVol	0.811 ***		0.812 ***		0.822 ***		0.706 ***	
UsDollarVol		0.811 ***		0.813 ***		0.822 ***		0.716 ***
Industry	0.021	0.022						
Governance	0.000	0.000	0.000	0.000				
Size	0.001	0.001	0.000	0.000	0.001	0.001		
No. of Obs.	104	104	104	104	104	104	104	104
Adjusted R^2	0.298	0.298	0.304	0.304	0.303	0.303	0.231	0.232