

# **TICK SIZE, MICROSTRUCTURE NOISE AND VOLATILITY INVERSION EFFECTS ON PRICE DISCOVERY IN OPTION MARKETS: THEORY AND EMPIRICAL EVIDENCE**

**Michal Czerwonko\***  
**Nabil Khoury\*\***  
**Stylianos Perrakis\*\*\***  
**Marko Savor\*\*\*\***

## **Abstract**

We document both theoretically and empirically a major dependence in both the Information Shares (IS) and Component Shares (CS) approaches to the estimation of the price discovery metrics on the errors arising out of the inversion method of the option value to find the implied stock price. We then introduce accurate inversion methods that result in a major increase in the information shares of option markets for both IS and CS metrics compared to the dominant Lagged Implied Volatility (LIV) inversion method. We apply these insights by examining the impact of the tick size reduction introduced by the CBOE in 2007 in its second pilot program on the simultaneous price discovery process in the markets for options and their underlying securities. In all cases we document a major impact of the tick size reduction in the option market that increases the option market information shares for all metrics and all inversion methods.

\*McGill University, Concordia University and University of Quebec in Montreal

\*\*University of Quebec in Montreal

\*\*\*Concordia University

\*\*\*\*University of Quebec in Montreal

# TICK SIZE, MICROSTRUCTURE NOISE AND VOLATILITY INVERSION EFFECTSON PRICE DISCOVERY IN OPTION MARKETS: THEORY AND EMPIRICAL EVIDENCE

## I. Introduction

On January 26, 2007 the CBOE introduced the first phase of its experimental *Penny Pilot Program* for 13 option classes that reduced the minimum tick size to \$0.01 for option series trading below \$3 and to \$0.05 for series trading above that level.<sup>1</sup> Similarly, the tick size for all option series on the QQQ, the DFX and the XSP indexes was also reduced to \$0.01. A second phase of the program was later introduced on September 28, 2007 and covered 22 additional option classes<sup>2</sup> which are the subject of this study. Similar tick size reductions have been introduced and studied earlier in other financial markets. In almost all published studies the focus was on the effect that these changes had on market liquidity. These studies concluded that tick size changes did have some impact on various liquidity measures such as bid-ask spread, trading volume and market depth, but that its size and significance were unclear.<sup>3</sup>

Tick size reductions are also expected to affect the price discovery process, since they have an obvious impact on the minimum size of price changes and the estimation of the error variances. Very few studies have dealt with this topic, and only one to our knowledge refers to option markets<sup>4</sup>. Yet, the effect of tick size reductions is expected to be particularly significant in option markets, given the fact that their liquidity is typically lower than that of corresponding equity markets.<sup>5</sup> A major complicating factor in such discovery studies is the error introduced by the option price inversion method in order to compute the underlying price.

In this paper we focus exclusively on the impact of the tick size and the volatility inversion methods on the price discovery process in the option and underlying asset markets. The main contribution of the study is to demonstrate the major impact that the option inversion methods have on the price discovery metrics in option markets. We show theoretically under plausible structural models of the inversion and microstructure errors, and with asset dynamics consistent with the available empirical techniques, that the errors introduced by the inversion methods reduce significantly these metrics. We develop test statistics that allow the assessment of the accuracy of an inversion technique under assumed asset dynamics of quite general type. We also present novel inversion methods that diminish the implied price error and allow for a more accurate estimation of the innovations that take place in the option market. In the empirical work, we show that, quite independently from the tick size reduction, these inversion methods result in

---

<sup>1</sup> CBOE Regulatory Circular RG07-09, *Penny Pilot Program*, January 18, 2007.

<sup>2</sup> CBOE Information Circular IC07-150 *SEC Approval of Penny Pilot Expansion*, September 27, 2007.

<sup>3</sup> See, for instance, Bacidore (1997), Bessembinder (2003), Bollen and Busse (2006), Jones and Lipson (2001), and other references discussed later on in the text.

<sup>4</sup> See Chen and Gau (2009) for both futures and options.

<sup>5</sup> The proportional option bid-ask spreads are typically more than twenty times as large as the equivalent spreads for the underlying assets, even for the most liquid index options; see also note 9.

major increases in the price discovery metrics of the option market as compared to the more traditional methods used in earlier studies. We also show that the tick size reduction results in a major increase in the option market discovery metrics for *all* inversion methods, the ones proposed in this paper as well as the traditional ones.

Price discovery refers to the process by which new information is embodied in trading activity. For optioned stocks such discovery can take place either in the option market or in the market for the underlying asset. It is known that there are advantages and disadvantages in using the option rather than the underlying asset for price discovery, but there are few studies of such a process. A theoretical study by Easley, O'Hara and Srinivas (EOS, 1998) develops conditions under which informed traders choose the option rather than the underlying asset market to exploit their informational advantages.<sup>6</sup> Nonetheless, their results refer mainly to option trading volume as a signal for informed trading, without any analysis of the simultaneous pricing in the two markets. Other theoretical studies are by Biais and Hillion (1994) and Capelle-Blancard (2003). However, most of the related literature is mainly empirical, and even here there are again very few studies involving option markets.

The price discovery process involving more than one market for the price of a given financial asset has been first analyzed in the pioneering work of Hasbrouck (1995), who developed the Information Shares (IS) approach.<sup>7</sup> Hasbrouck assumes that new information in each market is embodied in a common efficient or "true" asset price, together with independent error terms around that price when it is observed in the various markets. He then focuses on a particular representation of this efficient price known as the Vector Error Correction Model (VECM), originally developed by Engle and Granger (1987). IS are defined as the contributions of each market to the total variance matrix of the permanent component of the efficient price, evaluated from the VECM. An alternative metric of this permanent component, the Component Shares (CS), was developed by Gonzalo and Granger (1995). Most studies have used primarily IS in the study of simultaneous pricing of cross-listed securities in two different markets.<sup>8</sup> In our case it turns out that both measures yield virtually identical conclusions in all cases.

As already noted, there are very few studies of price discovery in the context of simultaneous trading in stock and option markets. In spite of the stylized theoretical analysis in EOS, there are objective factors that make the application of the IS and CS metrics difficult and imprecise. First of all, option markets are generally characterized by lower liquidity and wider bid-ask proportional spreads compared to their equity counterparts. While trading takes place within the quoted bid-ask spread, the effective spreads that result from observed trades are also very wide, thus rendering the concept of a unique implied stock price difficult to define.<sup>9</sup> Second and most important, the recovery of the underlying stock price from the observed option price is also

---

<sup>6</sup> The EOS model was applied empirically by Pan and Poteshman (2006).

<sup>7</sup> Earlier empirical studies had adopted the lead-lag analysis, which is a less robust methodology for this study. For an example of such a study see Stephen and Whaley (1990).

<sup>8</sup> See, for instance, Bacidore and Sofianos (2002), Eun and Sabherwal (2003) and Normandin (2004). For the relative merits of the two metrics see below, note 23.

<sup>9</sup> In the option sample used by Khoury, Perrakis and Savor (2011) the average option price was \$1.33 for the period 2002-2004, while the quoted and effective average bid-ask spreads never fell below 0.16 and 0.13 respectively, yielding respective proportional spreads of 12% and 10%. By contrast, in Fleming *et al* (1996) the average spread for stocks in their sample was 0.58%.

fraught with imprecision, since the accuracy of the estimation depends on the implied option model. Note that these difficulties in the implied price estimation are peculiar to the option markets and do not affect, for instance, the futures markets, for which the inversion of the cost-of-carry model is straightforward and does not involve unknown parameters or large errors.<sup>10</sup>

In an important empirical study of price discovery in option and underlying stock markets Chakravarty, Gulen and Mayhew (CGM, 2004), established a methodological approach for the estimation of the IS according to the Hasbrouck (1995) analysis. CGM use the option bid-ask spread midpoint as an estimate of the option price. They still need the implied volatility in order to estimate the option implied stock price. Since using the contemporaneous stock and option prices to find the implied volatility would result in a tautology, they invert the option price using a half hour lag to estimate the implied volatility, thus assuming that this volatility remains constant during that lag. With the time series of the observed stock prices and options' implied prices, CGM estimated the Hasbrouck IS metric and concluded that the option market's contribution to price discovery was about 17% on average. We shall call the CGM inversion method the Lagged Implied Volatility (LIV) method.

The main contribution of this paper is to demonstrate both theoretically and empirically that accuracy and efficiency of the implied price estimation method are fundamental for the estimation of the IS and CS metrics. We first show theoretically that any extraneous noise in this price will introduce a downward bias into these metrics.<sup>11</sup> Given the importance of the accuracy of inversion methods, we develop a statistical framework to *a priori* appraise such methods. In deriving this framework, we assume the stochastic volatility dynamics for the underlying price<sup>12</sup> without using any equilibrium arguments necessary for the derivation of the option price. Specifically, we define the volatility of the implied price and its correlation with the option price in accordance with the observed option market dynamics but *independently* of the inversion method. We then compare these two parameters with those of alternative inversion methods.

We demonstrate that the LIV inversion method as applied by CGM produces estimates of the option market's contributions to IS and CS that are as little as half their "true" measures. We provide clear evidence that the source of this bias lies in the extraneous noise or the measurement error of the parameters' estimation inherent in the LIV inversion method. For this reason we apply smoothing to the parameters used in those inversion methods, which efficiently aggregates the microstructure noise present in point-in-time estimates such as the LIV method. We test this smoothing with simulated data and find that it eliminates the error arising out of the inversion

---

<sup>10</sup> For instance, Constantinides *et al* (2011) report that for 1990-2002 the S&P 500 index futures basis risk error was less than 0.5% for 95% of the observations.

<sup>11</sup> This extraneous noise corresponds to the measurement error problem inherent in estimating the inversion parameters from the option price. This comes from the fact that the 'true' option price is roughly uniformly distributed on a relatively wide segment of the bid-ask spread while the 'true' underlying price is also distributed on some relatively narrow segment of its spread. The inversion methods presented in this work are largely concerned with those measurement errors.

<sup>12</sup> We assume the stochastic volatility framework to be sufficient for the objectives of this work since jumps either in the underlying security or in the volatility are rare events. Carr and Wu (2009) derive relatively low estimates for the error resulting from neglecting jumps in deriving the variance swap rate from the option prices.

method, while preserving this microstructure noise of the option market which should play a role in the price discovery.<sup>13</sup>

The smoothing we apply consists of using the median value for a set of the inversion parameters estimated within a lagged moving window relative to each time point of our implied price series.<sup>14</sup> We apply this smoothing technique to the LIV inversion method and call the result the Median LIV (MLIV). We also we develop and apply an entirely new inversion method based on the assumption of the homogeneity of the option price with respect to the underlying and strike prices, which is preserved under most stochastic volatility and jump-diffusion models, termed the Homogeneity Inversion (HI), whose parameters are also smoothed. It turns out that both methods yield remarkably similar estimates of the implied price and, more to the point, remarkably similar results with respect to the contribution of the option market according to the IS and CS metrics. As noted earlier, this contribution is more than twice the estimate from the conventional LIV method. This is true both before and after the change in tick size.

Turning now to the tick size reduction effect, we begin with the theoretical studies of its impact on the affected financial market. Seppi (1997), Anshuman and Kalay (1998) and Sandas (2001) presented stylized equilibrium models of competitive market making in which tick size and other trading restrictions play a role. Surprisingly, the tick size reduction does not always result in more liquid markets in these models. As Seppi (p. 104) noted, the relation between the tick size and liquidity is non-monotone and discontinuous. Similarly, Anshuman and Kalay (1998) showed that a positive tick size is a barrier to competition in market making, which results in positive expected profits for market makers.

Although there are several empirical studies on the impact of tick size reduction on price discovery in equity and futures markets<sup>15</sup>, only one of them, by Chen and Gau (2009), examined the role of options in this discovery. The authors use the CGM methodology for the inversion of the option price and examine the impact on the IS metric before and after the tick size reduction in the underlying Taiwan index tracking fund. They assess the impact of this reduction on the corresponding markets for the index tracking fund itself as well as for the markets of its options and futures. The authors report that the tick size reduction in the tracking fund's market significantly increased its IS at the expense of both the futures and options. They attribute this finding to the lower transaction costs in the tracking fund induced by the smaller tick size. In addition to the new inversion methods and their impact on price discovery, our study provides the following empirical contributions over and above those of Chen and Gau (2009). First, we provide a framework for an in depth analysis of the variability of information shares. Second, we demonstrate that there is a high variability over time in the contribution of the option market to the price discovery, which implies that only a systematic comparative analysis made possible by our data can capture the impact of the tick reduction, free of other factors influencing this

---

<sup>13</sup> Note that - ideally - an inversion method should eliminate or at least significantly decrease the error of such inversion from the microstructure noise in the past observations for the option price, while simultaneously preserving this error at the inverted present observation.

<sup>14</sup>Data smoothing by using a running median achieves two important goals for the problem at hand: first, it aggregates the measurement errors on the inversion parameters; second, it deals with the expected non-normality of those parameters as opposed to a running mean. See Bowman and Azzalini (1997).

<sup>15</sup> See Section III for more details.

price discovery. It also follows that the relatively large size of the sample we analyze allows drawing powerful statistical inferences. Third, our results show that the conclusions that may be drawn from studying individual stocks are significantly different from those that may be drawn from studying indices or exchange traded funds.

Other studies examining the impact of tick size reduction on price discovery in derivatives markets have dealt mostly with index futures. One such study examined the IS metric before and after tick size changes in the Chicago Mercantile Exchange (CME) Nasdaq-100 futures, both floor-traded and E-mini.<sup>16</sup> This tick size reduction increased significantly the efficiency of trading in both types of futures, but it also resulted in an increase in the IS metric for the E-mini futures at the expense of the floor-traded ones. By contrast, another microstructure study of similar instruments following tick size reduction reached exactly the opposite conclusions, namely that the tick size change has *reduced* market efficiency, by discouraging arbitrageurs who provide liquidity to the market.<sup>17</sup> These results reflect perhaps the ambiguity produced by the theoretical studies mentioned above. As noted earlier, such results are not directly related to our study, given the ease of inverting the futures expressions to estimate the implied price.

The remainder of the paper is organized as follows. In the next section we present the price discovery model, including the IS and CS metrics, as well as our analysis of the new option price inversion methods in the context of the stochastic volatility option pricing model. Section III then presents the data and section IV the corresponding empirical analysis, which includes the comparison of the inversion methods with respect to the information shares of the option market, the impact of the tick size reduction and various robustness checks of the results; this section also includes a comparison with the results of previous studies. Section V concludes the paper.

## II. Methodology

A fundamental issue in price discovery involving simultaneous pricing in stock and option markets is the inversion of the option price. In the CGM study the LIV inversion method for the option implied price was largely motivated by ‘tautology avoidance’, i.e. by the fact that no model may be estimated when the implied price is equal to the underlying price. In a variant of the Black-Scholes-Merton (BSM) model a lagged implied volatility was used to invert the present price of an option into its implied price. Since the BSM expression was used merely as a translation device, this approach is free of the option model misspecification error.<sup>18</sup> Nonetheless, there is a major concern regarding this inversion method due to the empirical properties of its critical parameter, the lagged IV. The “true” option price is unknown at any moment, and its value is generally proxied by the midpoint of the (extremely wide) option bid-ask spread, which midpoint price obviously includes microstructure noise, as does the contemporaneously observed underlying price. As a result, while the LIV method uses an unbiased IV estimate, the error variance of this quantity is expected to be high because of the aforementioned error components. This high error variance translates into noisy estimates for the implied price.

---

<sup>16</sup> See Kurov (2008).

<sup>17</sup> See Chen, Chou and Chang (2009).

<sup>18</sup> See Berkowitz (2010) for an asymptotic justification of the practitioners’ BSM model, whose variant the LIV inversion method explores.

We begin this section with a detailed presentation of the error arising from the inversion method of the option implied price. This is followed by a detailed description of the econometric methods used to estimate the IS and CS metrics and then by a theoretical analysis of the impact of the inversion error on these metrics. Next we present a the framework for gauging the accuracy of the inversion methods under the commonly assumed model of stochastic volatility asset dynamics in option pricing. The section concludes with the presentation of two new inversion methods and their comparison with LIV.

### 2.1 The impact of the inversion method on the IS Metrics

Consider high-frequency time series of a derivative  $f$  and its underlying security  $S$ . Hasbrouck (1995) and Gonzalo and Granger (1995) present econometric frameworks for the estimation of the contribution of each trading venue to the price discovery process for a single security traded in multiple venues. Since either framework requires that the prices in those venues be cointegrated, we can't directly estimate the price discovery from the observables. However, since the possibility of arbitrage closely links option and underlying security prices, the inversion of the option prices yields time series cointegrated with the underlying price with a known cointegrating vector. We denote this series by  $I$ :

$$I_t = h(f_t, \theta_t), \quad (1)$$

where  $h(\cdot)$  is a function yielding the implied underlying price and  $\theta$  is a set of parameters necessary for the inversion, assumed known exactly. Since this is clearly not a valid assumption, empirical studies such as CGM and others approximate  $I_t$  by setting  $I_t = \hat{h}(\cdot) = BSM_S^{-1}(f_t, IV_t)$ , with BSM and IV denoting the Black-Scholes-Merton price and implied volatility respectively, observed by necessity at some earlier time  $t-\tau$ ; this corresponds to an estimate  $\hat{\theta}_t$  of  $\theta_t$ . Since the effect of this estimation is a major new element in this paper, we discuss at length the impact of possible estimation errors on the price discovery process further on.

First we examine the estimation of the IS and CS metrics. For some well specified inverting function  $h(\cdot)$  we have the following order-one cointegrated price vector:

$$P_t = \begin{bmatrix} S_t \\ I_t \end{bmatrix} = \begin{bmatrix} E_t + \varepsilon_{S,t} \\ E_t + \varepsilon_{I,t} \end{bmatrix}. \quad (2)$$

In (2)  $E$  denotes the (unobservable) common efficient price,<sup>19</sup> which is assumed to follow a random walk:

---

<sup>19</sup> The order-one cointegration follows from the assumption that there is a single common trend, the efficient price. See Stock and Watson (1988).

$$E_t = E_{t-1} + v_t, \quad (3)$$

where  $E(v_t) = 0$ ,  $E(v_t^2) = \sigma_v^2$  and  $E(v_t, v_s) = 0$  for  $t \neq s$ . The error terms  $\varepsilon_{s,t}$  and  $\varepsilon_{l,t}$  represent transient microstructure effects, namely random variables with zero mean whose k-th order autocovariance matrices depend only on k. Then, under the Granger Representation Theorem we have the following Vector Error Correction Model (VECM) specification:

$$\Delta P_{t-l} = \alpha \beta' P_{t-1} + \sum_{i=1}^l A_i \Delta P_{t-i} + e_t, \quad (4)$$

where  $\alpha$  and  $\beta$  are  $[2 \times 1]$  matrices respectively containing adjustment parameters and cointegrating vectors, the  $[2 \times 2]$  matrices  $A_i$  describe the short-term dynamics of the process,  $l$  is the optimal number of lags. We also denote by  $\Omega$  the covariance of the serially-uncorrelated error terms  $e_t$ . The term  $\alpha \beta' P_{t-1}$  represents the equilibrium dynamics between the prices.

The IS metrics are derived by a decomposition of the random walk variance, which may be estimated even though the common efficient price is not observable:

$$\sigma_{RW}^2 = \psi \Omega \psi', \quad (5)$$

where  $\psi$  are the (identical) rows of a  $[2 \times 2]$  matrix  $\Psi$  given by<sup>20</sup>

$$\Psi = \beta_{\perp} \left( \alpha_{\perp} \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \sum_{i=1}^l A_i \right) \beta_{\perp} \right)^{-1} \alpha_{\perp}', \quad (6)$$

the  $[2 \times 1]$  matrices  $\alpha_{\perp}$  and  $\beta_{\perp}$  are non-trivial orthogonal complements to  $\alpha$  and  $\beta$  in (4) and the  $A_i$  matrices are as in (4).<sup>21</sup> The IS metric for a given market is defined<sup>22</sup> as the proportion of the random walk variance that is attributable to the innovations in that market. If the covariance matrix  $\Omega$  is diagonal, then we have a clean decomposition into contributions of each market to the total variance of the permanent component, with the contribution for the  $j$ -th market equal to

$\frac{\psi_j^2 \Omega_{jj}}{\sigma_{RW}^2}$ . This diagonal property will hold if the underlying random walk hypothesis is well

satisfied by the data, i.e. if there is little contemporaneous correlation between the residuals in (4). If the off-diagonal elements are non-zero, which is the prevailing case in empirical work, then we apply the Cholesky decomposition to the covariance matrix  $\Omega$  and derive the lower and upper bounds on IS since this quantity varies with each ordering of variables in the Cholesky decomposition.

---

<sup>20</sup> See Johansen (1990).

<sup>21</sup> See Johansen (1995) for the estimation methods of these orthogonal complements. Note also that by using (6), we avoid estimating the matrix  $\psi$  via an impulse-response function with *a priori* unknown convergence properties. In fact, for our data those functions may require an excessive number of iterations to achieve any reasonable convergence.

<sup>22</sup> See Hasbrouck (1995).



The Gonzalo and Granger (1995) decomposition into the permanent and transitory components starts with the same VECM model (3). Then we have the following decomposition of the price vector:

$$P_t = C_1 g_t + C_2 Z_t, \quad (7)$$

where  $g_t$  and  $Z_t$  respectively represent the permanent and transitory components,  $C_1$  and  $C_2$  are loading matrices, and  $g_t = \alpha'_\perp P_t$  and  $Z_t = \beta'_\perp P_t$  with  $\alpha_\perp$  and  $\beta$  as before. Note that here the efficient price and the integrated of order one permanent component  $g_t$  need not be random walks. Of interest to us is the [2x1] matrix  $\alpha_\perp$ . Under the additional assumption that  $Z_t$  does not Granger-cause<sup>23</sup>  $g_t$  this matrix may be identified up to a non-singular multiplication matrix. The interpretation of the permanent component  $g_t$  is that it is a weighted average (linear combination) of the observed prices with the component weights in  $\alpha_\perp$ . Booth *et al.* (1999) and Harris *et al.* (1995) suggest measuring price discovery by using CS normalized to 1. The interpretation of the CS metric is that the market, which reacts the least to the innovations in the other market, will display the highest relative weight in the permanent component. An attractive feature of the Gonzalo-Granger (1995) approach is that statistical hypotheses may be tested by a  $\chi^2(1)$  test, which we apply in our empirical work.

To assess the effect of the inversion errors in the option market metrics we need to provide some structure to the error terms  $\varepsilon_{S,t}$  and  $\varepsilon_{I,t}$  of the time series  $(S_t, I_t)$  and their relationship to the random walk. We use two alternative such structural models. In the first one there is no informed trading and the error terms are mutually independent random variables and independent of the random walk term  $v_t$ , with respective variances  $\sigma_S^2$  and  $\sigma_I^2$ . The second one is a modified version of a model presented in Hasbrouk (2007, Ch. 9), itself an adaptation of the Roll (1984) microstructure model, that recognizes the presence of informed trading and possible momentum effects. Our main result holds under both assumptions, but our detailed proof will be given under the simpler assumption that  $\varepsilon_{S,t}$  and  $\varepsilon_{I,t}$  are time-independent and uncorrelated random error terms; in such a case the *inversion error*  $\hat{\varepsilon}_{I,t} \equiv h f_t, \theta_t - h f_t, \hat{\theta}_t$  can be considered part of the term  $\varepsilon_{I,t}$ .

Under the Hasbrouk-Roll model the randomness of the microstructure noise affects the efficient price revisions  $v_t$ , which becomes now the sum of a pure random walk term  $u_t$  and a component driven by the direction of trade. Following Roll (1984), this direction is represented by the variable  $q_t$ , which determines both the location within the bid-ask spread, but also possibly reflects private information and affects, therefore, the efficient price. The two components of the time series  $(S_t, I_t)$  have each one its own revision proportional to the direction of trade. Analytically, while (3) still holds we now have:

$$v_t = u_t + \lambda w_t, \quad q_t = w_t, \quad S_t = E_t + c_S w_t, \quad I_t = E_t + c_I w_t + \hat{\varepsilon}_{I,t}. \quad (8)$$

---

<sup>23</sup> See Granger (1980).

In (8) the inversion error term  $\hat{\varepsilon}_t$ , with

$E(\hat{\varepsilon}_t) = 0$ ,  $Var(\hat{\varepsilon}_t) = \hat{\sigma}^2$ ,  $Cov(\hat{\varepsilon}_t, v_t) = 0$ ,  $Cov(\hat{\varepsilon}_t, \hat{\varepsilon}_{t-j}) = 0 \text{ } j \neq 0$ , cannot be incorporated into the implied price error. Nonetheless, for both models (2)-(3) and (2)-(8) we prove in the appendix the following result.

***Proposition 1:*** The higher the variance of the inversion error  $\hat{\varepsilon}_t$  the lower the information shares attributed to  $I_t$  in both models (2)-(3) and (2)-(8).

As noted earlier, the parameters of the structural model are not observable, and the price discovery effects have to be assessed through the IS and CS metrics evaluated from (4). To compare these metrics, we note that IS measures the impact of innovations on the random walk variance, whereas CS is concerned with the linear impact of the innovations. It appears that the IS metric is more popular in the market microstructure literature; however, in a recent study, Yan and Zivot (2010) recommended using both metrics because the IS metric is subject to microstructure noise from high frequency data. In addition, CS does not rely on the random walk assumption, which is usually not satisfied by the data.<sup>24</sup>

To estimate system (4), which forms the basis for the estimation of our metrics, we use the Johansen (1990) maximum likelihood approach. Gonzalo (1992) examined five alternative estimation methods for VECM and found that the empirical estimates may vary considerably across those methods in spite of the superconsistency of each method, with clearly superior properties for the Johansen (1990) method. Before proceeding with the estimates, however, we must deal with the error inversion issue, which in view of Proposition 1 needs to be addressed. The procedures used are explained in the following subsections.

## 2.2 Assessing the accuracy of the inversion methods

We evaluate the accuracy of the inversion methods applied in this paper under the assumption of stochastic volatility asset dynamics. As argued in the introduction, we assume that it is a sufficient framework for the analysis of the inversion problem. The state variables, that is, the underlying price  $S$  and its volatility  $\sigma$  vary according to the following bivariate diffusion:

$$\begin{aligned} dS_t &= \mu_t S_t dt + \sigma_t S_t dw_1 \\ d\sigma_t^2 &= \delta_t dt + \kappa_t dw_2 \\ \text{cov } dw_1, dw_2 &= \rho dt \end{aligned} \tag{9}$$

---

<sup>24</sup> There is an ongoing debate in the market microstructure and the related econometric literature about the relative merits of the Hasbrouck (1995) and Gonzalo and Granger (1995) approaches. As De Jong (2002) points out, CS for a given market ignores the variance of an innovation in this market while it measures the weight of this innovation in the increment of the efficient price. On the other hand, IS measures the share in the total *estimated* variance of the efficient price contributed by a given market.

where  $dw_1, dw_2$  are elementary Wiener processes and  $\mu_t$  is a function at most of the stock price, while the parameters  $\delta_t$  and  $\kappa_t$  are generally functions of  $\sigma_t$  but not of  $S_t$ .<sup>25</sup> By applying Ito's lemma to a derivative  $f_t$  resulting from (9) by an unspecified risk neutral valuation process and dividing both sides by the derivative price  $f_t$  we have:

$$df_t / f_t = f_t^{(S)} S_t \sigma_t dw_1 + f_t^{(\sigma)} \kappa_t dw_2 + \text{other terms} / f_t, \quad (10)$$

where we use only those terms that contribute to the quadratic and cross variations, the 'other terms' clearly do not contain either Wiener process, and with the superscripts in parentheses denoting partial derivatives. Those 'other terms' will vanish in all further derivations.

Now we state two lemmas proven in the Appendix that establish statistical benchmarks by which we assess the accuracy of the inversion methods applied in this paper. To empirically apply the first of these two lemmas, we assume in addition that under the risk neutral transformation of the stochastic volatility asset dynamics (9) the option price is linear homogeneous in  $(S_t, X)$ , where  $X$  denotes the strike price. This implies that the following relation holds:

$$f_t = f_t^{(S)} S_t + f_t^{(X)} X \quad (11)$$

### Lemma 1

Under the state dynamics (9), the homogeneity assumption and the range of values of our data, we have the following approximate relation for the volatility of the implied price  $\hat{\sigma}(I)$  as a function of the realized volatility of the option price returns  $\hat{\sigma} df / f$  :

$$\hat{\sigma} I \cong \tilde{\gamma} \hat{\sigma} df / f, \quad (12)$$

where for a given contract-day,  $\tilde{\gamma}$  represents the daily median of  $f / |f - f^{(X)} X|$ , with  $f^{(X)}$  estimated by (10) and we use the square root of the realized variance<sup>26</sup> as the estimate for the option return volatility.

Since (12) is an approximation, we cannot test formally for the differences between this benchmark and the observed volatility of the implied price derived by the inversion methods applied in this study; however, we consider it plausible that a 'good' inversion method will result in a tight spread around this benchmark quantity.

### Lemma 2

---

<sup>25</sup> It is possible to state a more general stochastic volatility process with respect to the functional dependence of the parameters of this process. The presented specification is consistent with the homogeneity assumption implied by the HI inversion method. See Bergman *at al.* (1996) for the general properties of the risk neutral stochastic process arising out of (9).

<sup>26</sup> To estimate the realized variance, the relevant literature recommends using longer time intervals than the one-second frequency in our data because of the microstructure noise present in high-frequency data (see Andersen *at al.* (2002) for a survey). In our estimate we use this one-second frequency since the microstructure noise clearly plays a role in our estimates of the implied price.

Under the state dynamics (9) and *without* restrictions on the parameters of the process, we have the following relation between the correlation of the implied price with the underlying price  $\rho_{IS}$  and the correlation of the derivative price with the underlying price  $\rho_{fS}$  :

$$\rho_{IS} \geq |\rho_{fS}| . \quad (13)$$

Since the sample counterparts of the quantities in (13) may be easily estimated from the data, Lemma 2 establishes a powerful statistical benchmark to assess the accuracy of the inversion methods. In the following subsections, as well as in Section IV, we use Lemma 2 in numerical results demonstrating the superiority of our proposed new inversion methods over LIV.

### 2.3 Volatility Inversion Methods

As noted in the beginning of this section, the key observation leading to an improvement over the LIV inversion method is the fact that the parameter set  $\theta$  as defined in equation (1) is subject to measurement error. This error is distinct from errors arising out of microstructure effects such as, for instance, the tick size in the option market.<sup>27</sup> We also saw that the error in measuring  $\theta$  reduces the price discovery metrics for that market.

Unlike microstructure noise, the parameter measurement error arising out of the volatility inversion does not belong to the discovery process and needs to be eliminated. The “optimal” parameter measurement is by polynomial fitting of the cross section of options on the same underlying asset at any given time with respect to the strike price and time to expiration;<sup>28</sup> we term this method the Regression Implied Volatility or RIV. Unfortunately, RIV cannot be applied uniformly to our data set, since the number of available options differs between stocks; further, the computational burden is prohibitively heavy if we want to do data smoothing. For this reason we examine this method only as a benchmark and discuss its performance further on in this section.

We deal with the measurement error problem by data smoothing,<sup>29</sup> which leads to canceling out or aggregation of the errors on  $\theta$ . In practice, the data smoothing we apply consists in using a running median for the inversion parameter set  $\theta$  rather than a single lagged observation. The use of the median is motivated by its ability to deal with the non-normality and with possible data errors (outliers) that may still be present even after the screening that we apply. The relative performance of the smoothing technique in reducing the volatility inversion error while leaving the microstructure error unaffected is examined in the next subsection.

In our first inversion method we apply our smoothing technique to LIV and term this method Median LIV, or MLIV. The second method we propose uses the property of homogeneity of the option price with respect to the underlying and the strike prices, which holds under most pricing models. We also show that our inversion methods compared to LIV admit shorter estimation

---

<sup>27</sup> We discuss the distinction between these two types of errors further on in this section.

<sup>28</sup> This was first applied by Dumas, Fleming and Whaley (1998). See also Christoffersen, Heston and Jacobs (2010) and Berkowitz (2010).

<sup>29</sup> See Bowman and Azzalini (1997).

windows, thereby allowing the incorporation of more variability in the unobservable state variables. At any time, we observe simultaneously a stock price  $S_t$  and a cross section of options with a common maturity but differing with respect to their strike prices. For any option in a given cross section (except for the two extreme strike prices), the two neighboring options may be used to estimate the partial derivatives,  $f_t^{(X)}$  and  $f_t^{(S)}$ . Once we have these estimates for the partials from past data, we may then easily invert (11) for the implied price  $I_t$ . We have from (11) for any  $\tau = t - M, \dots, t - 1$ :

$$f_\tau^{(X)} \cong \frac{1}{2} [f_\tau(X + \Delta X) - f_\tau(X - \Delta X)] / \Delta X$$

and

$$f_\tau^{(S)} = f_\tau - f_\tau^{(X)} X / S_\tau, \text{ or } \frac{1}{f_\tau^{(S)}} = \frac{S_\tau}{f_\tau - f_\tau^{(X)} X} \quad (14)$$

By inverting (9) and using (10), we have the following formula for the implied price:

$$I_t = \frac{f_t - f_t^{(X)} X}{f_t^{(S)}} \cong \tilde{\alpha}_t f_t + \tilde{\beta}_t X, \quad (15)$$

where for a given moving window length  $M$ ,  $\tilde{\alpha}_t$  and  $\tilde{\beta}_t$  are the respective medians of  $\frac{1}{f_\tau^{(S)}}$  and  $-\frac{f_\tau^{(X)}}{f_\tau^{(S)}}$  as derived by (14),  $\tau = t - M, \dots, t - 1$ . Note that the implied price under this inversion

approach has a meaningful interpretation as a portfolio composed of  $\tilde{\alpha}_t$  options and  $\tilde{\beta}_t X$  in the money market account. In empirical applications, where we use the average implied price derived from several options, this portfolio interpretation holds, in this case as a portfolio of options with unequal weights and the money market account.

The choice of the length of the moving window to estimate the inversion parameters involves balancing several opposing effects. For the main results of this paper we used a five minute window, with 300 past observations in one-second intervals to estimate the inversion parameters. Such a relatively short time interval that allows sufficient aggregation of the microstructure noise appears to be the optimal choice for the problem at hand.<sup>30</sup>

#### 2.4 The Relative Performance of the Volatility Inversion Methods

As noted, a volatility inversion method must minimize the parameter measurement error while leaving intact the microstructure error. In order to test whether the proposed inversion methods

---

<sup>30</sup> The risk of inducing a serial correlation in the error of the efficient price in equation (2) is small since by construction the time series of the inversion parameters is smooth. Further, the optimal number of lags in the VECM model (3) is low, with the length of the time interval to estimate the inversion parameters varying between 60 and 1800 seconds.

achieve this objective we use simulations that mimic more or less the characteristics of our data. In such simulations we first assume the parameters known and use an option pricing model yielding a closed form solution. In that model we introduce microstructure error arising out of the minimum tick size and examine the correlations between the underlying and implied prices. Next, we assume that the option price parameters are unknown and compute the implied price series according to one of the volatility inversion methods under consideration. A “satisfactory” inversion method should produce implied price series that satisfy Lemmas 1 and 2 and have similar correlations with the underlying price series as the ones arising out of the microstructure noise when the parameters are known.

For the simulations we choose the Heston (1993) model, which is the most popular stochastic volatility model in option pricing and has a closed form solution. In that model the stock price instantaneous mean is constant, while in (9)  $\delta_t$  is mean-reverting and  $\kappa_t$  is proportional to  $\sigma_t$ . The model is an “incomplete markets” one, which admits a closed form solution only if we assume that the market price of volatility risk is proportional to  $\sigma_t$ . We use the following parameters for the base case: initial volatility equal to its long-term mean of 0.3, volatility of volatility 0.4, speed of mean reversion 1, correlation of Wiener processes -0.6, riskless rate 4%, stock drift 8%, price of volatility risk -5% and initial underlying price 100.

For the simulations 300 paths for the stock were generated in one-second intervals for one trading day, which is some 31,000 seconds. For these paths 15 option prices valued according to the Heston (1993) model ('true prices') with strikes and maturities described in the tables below were generated. In the second step, the prices for tick rules of 0.05 (below \$3) and 0.1 (above \$3) were generated. For the tick-adjusted or quoted midpoint prices, this price at time  $t$  was set equal to the quoted midpoint price at  $t - 1$  unless the midpoint price was at least one tick away from the model price. In this latter case the quoted midpoint price moved by a minimal number of ticks so as to at least exceed (be below) the 'true' price for an upward (downward) movement in the model price.

In the above setup it is apparent that the true implied price is simply the underlying price, since there is no model or parameter uncertainty and the sole factor that may cause those prices to diverge is the presence of frictions in the option market; we assume that there are no frictions in the underlying market. The prices were inverted for HI, LIV and MLIV model with the interval or lag of 300 seconds. In addition, the RIV model in the spirit of Berkowitz (2010), and Christoffersen *et al* (2010) was applied. In that model, for the same lag as for LIV, we fit a least squares regression of the implied volatility (IV) against a constant, the time to maturity, the strike price, and the squares of the latter two variables; finally, the fitted value for IV at time  $t$  was used for the inversion. The results for the performance of the inversion methods are presented in Tables 2.1 and 2.2 below, with nine contracts used for HI inversion for each maturity since we need to discard two contracts, while all 15 contracts were used for the regression necessary for RIV inversion.

**[Table 2.1 approximately here]**

[Table 2.2 approximately here]

In Table 2.1 the benchmark (12) was adjusted to account for microstructure noise, since it was estimated by using the time  $t$  values of  $f/|f - f^{(X)}X|$  that contain that noise. The MLIV results are very close to the approximate benchmark (12) at the contract level in all cases, while the HI results are somewhat higher and the LIV and RIV results are much higher in all cases. On the other hand, at the aggregate level for all contracts the HI result is very close to the true volatility, while the MLIV is somewhat lower and the other two methods much higher. The results in Table 2.2 are even more interesting. First of all, we observe that in the absence of frictions the correlation between the derivative price and the underlying price is very close to 1 in all cases, as in the BSM model; this implies that the latter model, with its random walk assumption, may approximate quite well the more complex Heston model asset dynamics from the price discovery point of view. Second, the microstructure noise arising out of the minimum tick size lowers dramatically the size of the correlation  $\rho_{fs}$ , thus illustrating the importance of microstructure noise in the price discovery metrics. As for the correlation  $\rho_{fs}$  and its consistency with relation (13) of Lemma 2, we observe that the MLIV results satisfy it with equality in almost all cases, while the HI results are somewhat lower and the LIV and RIV much lower than  $\rho_{fs}$  in all cases; the latter results are also very similar to each other.

The results of Tables 2.1 and 2.2 remained qualitatively almost completely unchanged when the initial conditions were varied within the Heston model. Setting the initial volatility to 0.6 rather than 0.3 showed again that the MLIV method performed best and was very close to the benchmarks (12) and (13), while HI performed slightly worse and the other two methods much worse. We also varied the initial stock price to 300 and 1000 and again observed that the correlation  $\rho_{fs}$  performed as expected: it remained close to 1 in the absence of the tick size frictions, while it fell when the tick size was introduced; the drop was much lower at the price of 1000, reflecting the fact that the option prices were also much higher.<sup>31</sup> We conclude, therefore, that our moving window data smoothing achieved its objective of preserving the microstructure noise while eliminating most of the volatility inversion errors. Further, there is little or no improvement to be achieved by the polynomial fitting of the RIV method, which allows us to avoid the heavy computational burden of data smoothing in this case.

Figure 1 provides a graphic example of the implied prices derived by the HI and MLIV methods compared to the LIV method. The figure also plots the midpoints of the option and underlying prices. It is clear from the figure that the HI and MLIV implied prices change only because of the changes in the option price as we may expect *a priori*, while it is apparent that the implied

---

<sup>31</sup> The results are available from the authors on request.

price for the LIV method contains variability not justified by the option dynamics. For instance, we observe a large variability in the implied price for the LIV inversion for the flat segments in the option price; we also observe directional changes for this method when the changes in the call price conform to short lived demand/supply imbalances resulting in the widening or narrowing of the bid-ask spread shown in the graph as the call price bounces back and forth around the same level. These anomalies, which clearly reflect the option dynamics, don't appear in the implied prices derived by the HI or MLIV methods.

**[Fig 1 approximately here]**

### **III. Data**

Data for this study cover a 13 month period spanning from January 2007 to March 2008 inclusive, and includes tick by tick quotes on options and their underlying securities. More specifically, the main data set is comprised of 20 option classes out of 24 included in the second CBOE pilot program mentioned in the introduction and their corresponding underlying securities. Two option classes, namely Dendreon Corp. (ticker DNDN) and Financial Select Sector SPDR (ticker XLF), were dropped from the sample because a substantial number of days were missing in the CBOE tape. In addition, the Dow Jones (ticker DJX) and the mini S&P 500 (ticker XSP) option classes were also dropped because the underlying data is reported approximately every 10 to 15 seconds, which precludes applying the VECM model on a one-second interval. Moreover, as mentioned in the introduction, the assets underlying the last two option classes are non-tradable, whereas our study focuses only on options with publicly tradable underlying assets.

We included in our sample options maturing between 8 and 120 days. The data was first screened for standard arbitrage violations, excessive bid-ask spreads and excessive implied volatility.<sup>32</sup> Also, we screen out the contracts for which in a substantial proportion of quotes we observe  $c_A \geq S_A - X \geq c_B$  or  $p_A \geq X - S_B \geq p_B$ , since the option prices whose bid and ask prices bracket the early exercise value are not informative. Since in our final sample we included only close to ATM options, the above screening took place only for a limited number of close to maturity contracts. Note also that the presence of dividends has little effect in our data since the only optimal time to exercise a call option is at the end of the last *cum* dividend day.

---

<sup>32</sup>Our screens resulted in discarding less than 0.5% of the option quotes.



One-second time series were then constructed for options and their underlying securities. Gaps in the data were filled out with the last available quote. Finally, implied prices were derived for the three nearest to the money call or put options for each maturity within the above limits. The final restriction limited the moneyness of the included contracts, defined as the daily mean of the ratio of the strike to the underlying prices, to a range of 0.9 to 1.1. Altogether, the data includes 273 trading days or 13 calendar months. More specifically, for the period before (after) the policy change, the data includes 146 (127) trading days, 2940 (2520) company-days and 38673 (36470) option contracts. Table 3.1 presents the description of the underlying securities in our sample.

A comparison sample was also constructed using the same sampling approach. This comparison sample is composed of the same number of randomly selected option classes where all available classes in the Option Metrics database were first filtered according to the following criteria: i. the total volume of transactions of the option class is no higher than that of the most traded class in the main sample and no lower than that of the least traded, ii. the class remains actively traded during the entire period under study.

**[Table 3.1 approximately here]**

Table 3.2 presents the descriptive statistics of the main and control samples for the 7 months before and the 6 months after the minimum tick size reduction was implemented. The table shows that both quoted and effective spreads declined in a statistically significant manner after the minimum tick size reduction was implemented. It is important to note that prior to the change the differences in spreads between the main and control samples were not statistically significant either for quoted or for effective spreads. These last results provide further evidence of the effect of the minimum tick size reduction on transaction costs in option markets. Overall, these findings are consistent with those reported for equity markets by, among others, Henker and Martens (2005), Bessembinder (2003), Ronen and Weaver (2001), Foucault (1999), and Bollen and Whaley (1998) and for futures markets by Kurov (2008) who found that the reduction in the minimum tick size reduced the effective spreads for the floor traded and the electronically traded E-mini Nasdaq 100 index futures<sup>33</sup>.

**[Table 3.2 approximately here]**

---

<sup>33</sup> On the other hand, Bollen and Busse (2006), Chakravarty et al (2004), Graham et al (2003), Jones and Lipson (2001) and Harris (1994) suggest that the reduction in tick size can increase the cost for liquidity providers thus adversely affecting the bid-ask spread.<sup>3</sup>

The results in table 3.2 also indicate that the reductions in spreads were accompanied by corresponding reductions in the sizes of the quotes. More specifically, the table shows that both the bid and ask quote sizes of the option classes included in the pilot project decreased significantly after the minimum tick size reduction was implemented, while the corresponding sizes increased moderately and non-significantly for the option classes in the control sample. This preliminary result is consistent with the findings reported by Bacidore (1997) for the Toronto Stock Exchange subsequent to decimalization.

These introductory results provide a preliminary indication of the importance of the tick size change on transaction costs and the efficiency of trading. In the traditional analyses such liquidity and efficiency changes should not have affected the price discovery process in the markets for options and their underlying securities, given that price discovery is assessed at the midpoint of the bid-ask spread. In our case, though, these efficiency changes reduce the errors of the option inversion process and, thus, are expected to improve the information share metrics. However, before turning to this issue, it is also interesting to note in table 3.2 that the pilot and non-pilot samples exhibit statistically significant differences in price and volume levels, while their volatility levels are approximately equal prior to the tick size change. All three variables are significantly different between the samples after the tick change. Similarly, price, volume and volatility levels all increased after the tick size reduction regardless of the pilot project, reflecting the general market environment at the time of the pilot project's implementation.

#### **IV. Empirical analysis**

To derive our main results we use a five-minute moving window (or lag) of data for the estimation of the parameters of the inversion methods. Results for other time-frames that we used are presented as robustness checks.

##### 4.1 Accuracy of Inversion Methods

In this section we present the accuracy of the inversion methods relative to the two benchmarks (12) and (13). Apart from assessing the merits of the two methods proposed in this paper relative to LIV, the empirical results are of some interest since to our best knowledge our study is the first to assess the implied price *a priori*. We focus our presentation on the standard errors of the observed quantities; specifically, we present the descriptive statistics for the following standardized differences from the benchmarks, or *t*-stats:

$$t = \frac{Q - B}{SE(B)} , \tag{16}$$

where  $Q$  is a given quantity for the implied price,  $B$  is the benchmark for this quantity and  $SE(B)$  is the standard error for a given benchmark.<sup>34</sup> We derive the results of the implied prices for all the option contracts used in our study separately for the periods before and after the minimal tick policy change<sup>35</sup>.

Table 4.1 displays descriptive statistics for the correlation between the implied and underlying price. These results show that the MLIV and HI methods provide a viable alternative to the standard inversion using lagged implied volatilities. The correlations are presented for the various inversion approaches applied both before and after the change in minimum tick size. The primary results show that the LIV method performs poorly compared to MLIV and HI because it clearly underestimates the correlation between implied and actual equity prices. Moreover, the LIV inversion method clearly biases downward this correlation, with a large proportion of option contracts with implied prices having a correlation significantly below the benchmark (13). For HI and MLIV methods the data falls in a tight interval around the benchmark in both periods, whereas the performance worsens in the second period for LIV.

**[Table 4.1 approximately here]**

Another interesting result in Table 4.1 is the important increase in the standardized correlation between actual and implied underlying security prices after the tick reduction was implemented. This result can be observed irrespective of the inversion method under study, both when the implied prices are aggregated on a contract level basis as presented in Table 4.1, but also by underlying security in unreported results.<sup>36</sup> This increase in correlation is clearly a result of the tightening of the bid-ask spreads for options following the tick size reduction as in Table 3.2, since tighter spreads allowed option market quotes to more closely mimic the dynamics of their corresponding underlying securities, thus leading to a greater correlation between the two markets.

**[Table 4.2 approximately here]**

Table 4.2 presents descriptive statistics for raw correlations aggregated at the security-day level.<sup>37</sup> The results in this table confirm the findings in Table 4.1. Indeed, the correlation for LIV is, on average, approximately half that for HI and MLIV. We again observe significant

---

<sup>34</sup>For the correlation between the implied and underlying price and its benchmark (13), we applied the Fisher transformation, which assures approximate normality.

<sup>35</sup>To facilitate the interpretation of the results note that a median of, say, -1.70 means that for 50% of the option contracts the implied price yields a figure lower than its benchmark at the 5% or lower significance level.

<sup>36</sup>All unreported results are available from the authors upon request.

<sup>37</sup>Descriptive statistics for raw correlations at the contract level are not presented here since they are not qualitatively different from those in Table 4.2. The raw correlations for the aggregate implied price were generally higher than the average correlation at the option-contract level.

increases in the average or median correlations after the policy change, that are also confirmed in the results for the raw correlations at the contract level. This finding underscores the role of the tick size in the observable comovement of the two markets.

**[Table 4.3 approximately here]**

The results presented in Table 4.3 show the standardized volatility of implied price relative to the benchmark (12). To facilitate the interpretation of the results, note that ten standard errors correspond to approximately 4.8% of the benchmark volatility (12).<sup>38</sup> Even if we take into account that this benchmark is weaker than the equivalent result (13) for the correlation coefficient, it is clear that the LIV method results in a noisy implied price compared to HI and MLIV. More specifically, the results show that the LIV is more than double, on average, the corresponding benchmark quantity. The equivalent unreported results for implied prices aggregated at the underlying security level also show that the LIV approach presents the highest volatility of implied price relative to the underlying price irrespective of the inversion approach. For the latter two methods, the volatility is tightly distributed around the benchmark (14) while for the former method it is clearly distributed above this benchmark indicating extraneous noise in the implied price.

To sum up, the results in Tables 4.1 to 4.3 clearly indicate that the LIV inversion method underperforms when compared to the two methods proposed in this paper in terms of the basic time-series properties of the implied price. Similar results for moving windows or lags up to thirty minutes were not qualitatively different and are not presented here.

#### 4.2 Option markets information shares

In this section we present our main results for the IS and CS metrics, which assess the price discovery in the option market before and after the minimal tick policy change. In addition to proposing an improved methodology for inverting the implied equity price from option markets, one of the main objectives of the study is to analyse the price discovery performance of option markets in comparison with their equity counterparts. The paper thus studies the effect of the alternative inversion approaches on the importance of option markets in the price discovery process and challenges some of the traditional inferences that have been drawn using the typical LIV method. Since we also observe rather substantial variability in these metrics over time in general, we contrast the results for our main sample with a comparison sample.

Table 4.4 presents the summary results for the IS metric. It is clear that the average IS increased after the tick size reduction irrespectively of the inversion method for both main and comparison samples. While this reflects the state of the pair of financial markets between the two periods, the

---

<sup>38</sup> In other words, 210 standard errors correspond to twice the benchmark (14).

differential increase in the two samples between periods is much higher for the main than for the comparison sample; this difference is also statistically significant. More to the point, we note that all IS metrics are much higher for the HI and MLIV inversions compared to the LIV inversion, For the IS midpoint the LIV inversion yields about half the IS values before the policy change, and only about 36% of the IS values of the other two methods after the tick size reduction. We attribute this difference to the noisiness of the LIV method documented above. We conclude that the LIV method leads to an underestimation of the role of the option markets in price discovery. Finally, we observe highly consistent results between the HI and MLIV inversions.

**[Table 4.4 approximately here]**

Figure 2 plots the daily averages across underlying securities for the midpoint IS metric for the main and comparison samples under the HI inversion method.<sup>39</sup> We observe a significant increase in IS for the main sample after the tick size reduction (day 0 in the figure) without a corresponding increase for the comparison sample. This figure also underscores the importance of the comparison sample since we observe several large increases and decreases in IS over time. Note also an apparent isomorphism between the two plots before the policy change, which may imply market-wide factors driving the price discovery in the option markets.

**[Fig 2 approximately here]**

The results in Table 4.5 for the CS metric confirm the findings in Table 4.4. The results demonstrate the importance of adequately dealing with the methodological bias that is present in the traditional inversion approach and show that the proposed methods' improvements result in a much more important role of option markets in price discovery than would be traditionally expected. In addition, the proportion of component shares CS that are statistically significant at a 5% level significantly increases after the tick size reduction. Above all, both IS and CS metrics for option markets have increased quite significantly after the reduction of the minimum tick size irrespective of the inversion method. This effect appears even more clearly in the context of the improved MLIV and HI approaches.

**[Table 4.5 approximately here]**

Table 4.6 presents the results for the minima and maxima of the IS metric aggregated for each underlying separately for the periods before and after the policy change for the HI and LIV inversion methods.<sup>40</sup> The results in this table confirm previous findings both with respect to the effect of the policy change and to the impact of the inversion method on this metric. This table also shows that the option market information shares are much higher for the four exchange traded funds (ETF) in our sample than for equities, differing by as much as a factor of ten.

**[Table 4.6 approximately here]**

---

<sup>39</sup>The MLIV inversion yielded similar results.

<sup>40</sup>The results for the MLIV inversion method were not qualitatively different from those of the HI method.

As noted in the introduction, there are relatively few studies of the price discovery process in underlying and option market pairs. CGM (2004) studied this process for 60 optioned firms' equities over the five-year period 1988-1992. The average IS for the option market was about 17% and there was comparatively little dispersion among firms. These results are far apart from ours for the equities, since the LIV inversion method in Table 4.6 yields IS measures before the tick size change that are about one tenth of the CGM numbers, and even the HI inversion method's numbers are also significantly lower than in CGM. After the tick size policy change the option market's IS metrics become generally higher for our firms than in CGM for HI but not necessarily for LIV. Note that our main sample has very little commonality with that of CGM, since not only is the time period different by more than fifteen years, but also only two firms, BMY and GM, are common to both samples.

Our IS results in Table 4.6 for the options of the four ETFs are also lower than those of Chen and Gau (2009) for the options on the Taiwan index-tracking fund for the LIV inversion method, but the difference is not as large as in the comparison with CGM for equities. The HI inversion method makes the options' IS metric comparable or higher than that of Chen and Gau. Note that in this latter study the tick size decreased for the ETF but not for the option market, with the result that IS decreased for the latter market after the policy change.

#### 4.3 Robustness Checks

Table 4.7 presents the results for 15- and 30-minute moving windows (or lags) for the estimation of the inversion parameters. To shorten the presentation we present the IS metrics' results only for the midpoint. The Table shows that our results are robust to the length of the window (or lag).

**[Table 4.7 approximately here]**

Several papers on market microstructure such as Hasbrouck (1995) or Chakravarty *et al.* (2004) use a polynomial distributed lag model (PDL) to estimate the VECM model. The main role of PDL is to reduce the number of the estimated coefficients in case a large number of lags is applied since there is no good econometric justification for this approach other than the precision of the estimation. We did not use PDL to derive our main results since we applied a relatively low number of lags indicated by the Schwartz information criterion. As a robustness check we verify if the inclusion of a large number of lags may affect our results. The outcome of this analysis is summarized in Table 4.8. This table shows that the IS and CS metrics are robust to the inclusion of a large number of lags for the HI and MLIV inversion methods. For the LIV method, these metrics increase when the included number of lags increases in the period before the policy change. For all inversion methods, the proportion of statistically significant CS decreases when the included number of lags increases. We conclude that our results are robust to the inclusion of a large number of lags in the estimation of the VECM model.

**[Table 4.8 approximately here]**

## V. Conclusions

We present an alternative approach to the price discovery process in option markets, in which we stress the importance of the option price inversion method in order to estimate the implied underlying price. We show theoretically that the price discovery metrics are highly sensitive to the errors introduced by the inversion method, and we develop theoretical benchmark measures in order to assess the inversion methods under plausible underlying asset dynamics. We also verify empirically with simulated data and a known option pricing model the adequacy of our proposed methods in dealing with the inversion errors. We also show that the dominant lagged implied volatility inversion method, which was used exclusively in earlier studies, tends to underestimate seriously the contributions of the option markets to price discovery.

We apply these insights to estimate the price discovery effects of a change in minimum tick size in the option market implemented recently in a pilot study of 22 optioned stocks in the CBOE. We examine price discovery by using this pilot study sample and a matched sample of optioned stocks in which the minimum tick size was not changed. We show that tick size change had a major impact on the information shares of the option market, increasing the importance of this market in the discovery process. This increase was present for all option price inversion methods, but it was much more pronounced for the new inversion methods that we introduced. It was also much more pronounced for the options on exchange-traded funds than on individual stocks.

Table 2.1: Performance of the inversion methods for the implied price volatility in the Heston (1993) model (Monte Carlo averages)

T	K/S0	True (vol S)	Benchmark (eq. 12)	HI	LIV	MLIV	RIV
A: Contract level implied price volatility							
1/12	1.05	0.0112	0.0095	0.0117	0.0182	0.0105	0.0212
1/12	1.00	0.0112	0.0103	0.0123	0.0185	0.0108	0.0186
1/12	0.95	0.0112	0.0106	0.0122	0.0185	0.0108	0.0164
2/12	1.05	0.0112	0.0096	0.0116	0.0181	0.0104	0.0168
2/12	1.00	0.0112	0.0100	0.0126	0.0184	0.0107	0.0164
2/12	0.95	0.0112	0.0104	0.0128	0.0185	0.0108	0.0168
3/12	1.05	0.0112	0.0096	0.0129	0.0183	0.0107	0.0167
3/12	1.00	0.0112	0.0099	0.0125	0.0183	0.0106	0.0161
3/12	0.95	0.0112	0.0102	0.0125	0.0184	0.0107	0.0174
A: Aggregate implied price volatility							
n/a	n/a	0.0112	n/a	0.0094	0.0142	0.0084	0.0149

Parameter values are: initial volatility equal to its long-term mean of 0.3, volatility of volatility 0.4, speed of mean reversion 1, correlation of Wiener processes -0.6, riskless rate 4%, stock drift 8%, price of volatility risk -5% and initial underlying price 100.

Table 2.2: Performance of the inversion methods for the correlation between option implied and underlying prices in the Heston (1993) model (Monte Carlo averages)

T	K/S0	f-S w/o noise	f-S with noise	HI	LIV	MLIV	RIV
1/12	1.05	0.983	0.052	0.043	0.030	0.052	0.025
1/12	1	0.993	0.046	0.041	0.027	0.046	0.027
1/12	0.95	0.998	0.065	0.056	0.038	0.064	0.042
2/12	1.05	0.976	0.063	0.052	0.036	0.063	0.039
2/12	1	0.988	0.046	0.042	0.027	0.046	0.030
2/12	0.95	0.995	0.061	0.055	0.035	0.060	0.039
3/12	1.05	0.971	0.034	0.029	0.020	0.034	0.022
3/12	1	0.985	0.047	0.043	0.027	0.046	0.031
3/12	0.95	0.992	0.058	0.051	0.034	0.058	0.036
A: Aggregate implied price							
n/a	n/a	n/a	n/a	0.119	0.046	0.131	0.047

Parameter values are: initial volatility equal to its long-term mean of 0.3, volatility of volatility 0.4, speed of mean reversion 1, correlation of Wiener processes -0.6, riskless rate 4%, stock drift 8%, price of volatility risk -5% and initial underlying price 100.



Table 3.1: Underlying Securities in Main Sample

The table provides the name and ticker symbol for the underlying securities included in the main sample and includes all securities that were part of the tick size reduction pilot project.

<b><u>Ticker</u></b>	<b><u>Underlying Name</u></b>
AAPL	Apple
AMGN	Amgen
AMZN	Amazon
BMJ	Bristol-Myers Squibb
C	Citigroup
COP	ConocoPhillips
CSCO	Cisco
FCX	Freeport-McMoran
GM	General Motors
MO	Altria Group
MOT	Motorola
NYX	NYSE Euronext
QCOM	Qualcomm
RIMM	Research in Motion
T	AT&T
YHOO	Yahoo
DIA*	Diamonds Trust
OIH*	Oil Services HLDRS
SPY*	SPDR S&P 500
XLE*	Energy Sector SPDR

\* Exchange traded funds

Table 3.2 – Descriptive statistics before and after the tick size reduction

The table provides descriptive statistics for the securities included in the main sample. More specifically, the table provides means on a per security basis for *Price*, the national best offer and ask quoted prices, *Implied Volatility*, the Black-Scholes-Merton volatility implied in option prices, *volume*, the total daily volume, *Effective spreads*, twice the absolute value of the difference between transaction prices and the average between quoted bid and ask prices, *Quoted spreads*, the difference between quoted bid and ask prices, *Bid size* and *ask size*, the number of contracts available at the quoted bid and ask prices. The means are provided for both the main sample, securities included in the pilot project, and the control sample, securities not included in the pilot project. The t-statistics relate to the difference of means statistical test for the main sample.

	Before Policy Change			After Policy Change		
	<u>Main</u>	<u>Control</u>	<u>t-stat</u>	<u>Main</u>	<u>Control</u>	<u>t-stat</u>
<b>Price</b>	3.86	4.43	-71.59***	4.63	4.97	-38.92***
<b>Implied volatility (%)</b>	32.55	32.37	12.32***	42.52	45.80	-164.63***
<b>Volume ('000s)</b>	7.49	1.46	8.27***	10.05	1.71	898.75***
<b>Effective Spread</b>	0.128	0.171	-1.53	0.056	0.205	-355.77***
<b>Quoted Spread</b>	0.119	0.138	-0.67	0.062	0.166	-37.54***
<b>Bid size ('000s)</b>	1.87	0.45	652.02***	0.45	0.50	-54.71***
<b>Ask size ('000s)</b>	1.71	0.40	663.56***	0.44	0.48	-34.71***
<b>Number of observations</b>	3,567,193	814,125		5,306,591	910,475	

\*\*\*, \*\* & \* : indicates a statistically significant difference of means before and after the minimum tick size change at a 1%, 5%, 10% level respectively.

Table 4.1: Relative standard errors of correlations between actual and implied underlying security prices before (pre) and after (post) the tick size reduction (aggregated by contract)

The table provides descriptive statistics for relative standard errors of correlations between actual and implied underlying security prices before (pre) and after (post) the tick size reduction. More specifically, the errors are described by providing their *mean*, *median*, *sigma*, the standard deviation, *qrang*, *p1*, the 1<sup>st</sup> percentile, and *p99*, the 99<sup>th</sup> percentile. The information is aggregated on a per contract basis and is presented for various inversion methods.

<b><u>Inversion method</u></b>	<b><u>mean</u></b>	<b><u>median</u></b>	<b><u>sigma</u></b>	<b><u>qrang</u></b>	<b><u>p1</u></b>	<b><u>p99</u></b>
<u>Pre</u>						
HI	-0.103	-0.054	0.523	0.215	-0.886	0.393
LIV	-1.080	-0.942	1.162	1.411	-4.565	1.152
MLIV	-0.014	-0.011	0.095	0.084	-0.277	0.224
<u>Post</u>						
HI	-0.278	-0.205	0.486	0.378	-1.511	0.481
lag LIV	-5.577	-4.917	3.440	4.439	-15.749	-0.052
MLIV	-0.059	-0.051	0.233	0.142	-0.757	0.578

Table 4.2: Correlations between actual and implied underlying security prices before and after the tick size reduction (aggregated by underlying security)

The table provides descriptive statistics for correlations between actual and implied underlying security prices before (pre) and after (post) the tick size reduction. More specifically, the correlations are described by providing their *mean*, *median*, *sigma*, the standard deviation, *qrang*, *p1*, the 1<sup>st</sup> percentile, and *p99*, the 99<sup>th</sup> percentile. The information is aggregated on a per security basis and is presented for various inversion methods.

<b><u>Inversion method</u></b>	<b><u>mean</u></b>	<b><u>median</u></b>	<b><u>sigma</u></b>	<b><u>qrang</u></b>	<b><u>p1</u></b>	<b><u>p99</u></b>
<u>Pre</u>						
HI	0.084	0.069	0.067	0.088	-0.010	0.285
LIV	0.046	0.039	0.037	0.049	-0.010	0.161
MLIV	0.086	0.071	0.068	0.090	-0.009	0.288
<u>Post</u>						
HI	0.333	0.325	0.123	0.166	0.067	0.617
LIV	0.171	0.161	0.079	0.100	0.015	0.375
MLIV	0.336	0.328	0.123	0.166	0.069	0.620

Table 4.3: Standardized volatility of implied prices before and after the tick size reduction (aggregated by contract)

The table provides descriptive statistics for standardized volatility of implied prices before (pre) and after (post) the tick size reduction. More specifically, the standardized volatilities are described by providing their *mean*, *median*, *sigma*, the standard deviation, *qrange*, *p1*, the 1<sup>st</sup> percentile, and *p99*, the 99<sup>th</sup> percentile. The information is aggregated on a per contract basis and is presented for various inversion methods.

<b><u>Inversion</u></b>	<b><u>mean</u></b>	<b><u>median</u></b>	<b><u>sigma</u></b>	<b><u>qrange</u></b>	<b><u>p1</u></b>	<b><u>p99</u></b>
<u>Pre</u>						
HI	5.482	4.090	31.708	5.980	-8.955	27.951
LIV	249.074	108.353	7390.187	39.249	19.585	294.072
MLIV	9.332	2.761	447.559	22.818	-50.499	73.522
<u>Post</u>						
HI	4.767	3.090	16.432	5.386	-8.070	29.154
LIV	183.369	120.917	4837.497	48.070	33.814	317.749
MLIV	0.561	2.078	19.382	25.298	-44.941	43.365

Table 4.4: Average IS before and after the tick size reduction

The table provides descriptive statistics for information shares, *IS*, before and after the policy change reducing the minimum tick size. More specifically, the IS are described by providing their *min*, the minimum average value, *mid*, the median average value, *max*, the maximum average value. The information is provided for both the main sample, securities included in the pilot project, and the control sample, securities not included in the pilot project. The statistics are also presented for various inversion methods.

<b><u>Inversion Method</u></b>	<b><u>Main Sample</u></b>			<b><u>Comparison Sample</u></b>		
	<b><u>ISmin</u></b>	<b><u>ISmid</u></b>	<b><u>ISmax</u></b>	<b><u>ISmin</u></b>	<b><u>ISmid</u></b>	<b><u>ISmax</u></b>
<u>A: Before Policy Change</u>						
HI	0.043	0.062	0.081	0.014	0.025	0.036
LIV	0.025	0.032	0.040	0.012	0.016	0.020
MLIV	0.045	0.065	0.086	0.014	0.026	0.037
<u>B: After Policy Change</u>						
HI	0.169	0.321	0.473	0.038	0.078	0.119
LIV	0.058	0.118	0.179	0.023	0.035	0.047
MLIV	0.176	0.331	0.487	0.039	0.082	0.124

Table 4.5: Average CS before and after the tick size reduction

The table provides descriptive statistics for component shares, CS, before and after the policy change reducing the minimum tick size. More specifically, the CS are described by providing their CS, the average value, *pr.p*, the proportion of component shares that are statistically significant at a 5% level. The information is provided for both the main sample, securities included in the pilot project, and the control sample, securities not included in the pilot project. The statistics are also presented for various inversion methods.

<b>Inversion Method</b>	<b>Main Sample</b>		<b>Control Sample</b>	
	<b><u>CS</u></b>	<b><u>pr.p</u></b>	<b><u>CS</u></b>	<b><u>pr.p</u></b>
	A: Before Policy Change			
HI	0.146	0.432	0.112	0.394
LIV	0.072	0.372	0.055	0.304
MLIV	0.151	0.437	0.113	0.397
	B: After Policy Change			
HI	0.456	0.938	0.209	0.719
LIV	0.164	0.795	0.085	0.478
MLIV	0.469	0.943	0.217	0.734

Table 4.6: Average Information Shares per Underlying before and after Policy Change

The table provides descriptive statistics for information shares, *IS*, before and after the policy change reducing the minimum tick size. More specifically, the IS are described by providing their *min*, the minimum average value, *max*, the maximum average value. The information is provided for each security included in the main sample, securities included in the pilot project, and is presented for various inversion methods.

Ticker	Before Policy Change				After Policy Change			
	HI Inversion		LIV Inversion		HI Inversion		LIV Inversion	
	<u>Min</u>	<u>max</u>	<u>min</u>	<u>max</u>	<u>min</u>	<u>max</u>	<u>min</u>	<u>max</u>
AAPL	0.005	0.042	0.005	0.017	0.184	0.683	0.059	0.321
AMGN	0.005	0.020	0.006	0.012	0.064	0.329	0.021	0.090
AMZN	0.003	0.022	0.003	0.010	0.165	0.549	0.056	0.242
BMJ	0.010	0.025	0.012	0.019	0.100	0.350	0.034	0.120
C	0.008	0.026	0.009	0.016	0.095	0.309	0.030	0.100
COP	0.012	0.055	0.010	0.026	0.116	0.452	0.027	0.121
CSCO	0.008	0.018	0.007	0.012	0.126	0.468	0.045	0.196
FCX	0.010	0.041	0.011	0.020	0.181	0.460	0.043	0.130
GM	0.009	0.030	0.008	0.017	0.142	0.452	0.051	0.173
MO	0.017	0.043	0.009	0.020	0.102	0.337	0.032	0.101
MOT	0.010	0.025	0.009	0.016	0.135	0.408	0.052	0.161
NYX	0.012	0.039	0.009	0.017	0.117	0.394	0.032	0.111
QCOM	0.005	0.023	0.005	0.011	0.116	0.530	0.035	0.202
RIMM	0.004	0.025	0.003	0.009	0.130	0.602	0.033	0.220
T	0.011	0.031	0.007	0.016	0.138	0.483	0.043	0.182
YHOO	0.004	0.010	0.004	0.008	0.078	0.331	0.032	0.131
DIA*	0.127	0.206	0.074	0.108	0.258	0.458	0.095	0.174
OIH*	0.302	0.466	0.139	0.218	0.525	0.777	0.213	0.368
SPY*	0.072	0.157	0.035	0.068	0.210	0.503	0.077	0.213
XLE*	0.228	0.319	0.127	0.168	0.404	0.589	0.149	0.219

\* exchange traded funds

Table 4.7: Results for Varying Moving Window or Lag

The table provides average information shares, *IS*, and component shares, *CS*, both before (pre) and after (post) the policy change reducing the minimum tick size. The information is presented for various moving windows (or lags) for the estimation of the inversion parameters. The table also provides the proportion of CS estimates that are statistically significant at a 5% level. The statistics are provided for the main sample, securities included in the pilot project and for various inversion methods.

<b>Inversion Method</b>	<b>Main Results</b>		<b>10 minutes</b>		<b>30 minutes</b>	
	<b><u>pre</u></b>	<b><u>post</u></b>	<b><u>pre</u></b>	<b><u>post</u></b>	<b><u>pre</u></b>	<b><u>post</u></b>
	<u>Information Shares (midpoint)</u>					
HI	0.062	0.321	0.068	0.326	0.069	0.327
LIV	0.032	0.118	0.034	0.115	0.032	0.109
MLIV	0.065	0.331	0.071	0.335	0.071	0.335
	<u>Component Shares</u>					
HI	0.146	0.456	0.154	0.459	0.157	0.458
LIV	0.072	0.164	0.073	0.160	0.070	0.150
MLIV	0.151	0.469	0.159	0.472	0.159	0.471
	<u>Proportion p-value &lt; 5% for Component Shares</u>					
HI	0.432	0.938	0.397	0.924	0.356	0.894
LIV	0.372	0.795	0.356	0.754	0.327	0.715
MLIV	0.437	0.943	0.410	0.930	0.373	0.917

Table 4.8: Results for PDL Model

The table provides average information shares, *IS*, and component shares, *CS*, both before (pre) and after (post) the policy change reducing the minimum tick size. The information is presented for the polynomial distributed lag model (PDL) to estimate the VECM model. The table also provides the proportion of CS estimates that are statistically significant at a 5% level. The statistics are provided for the main sample, securities included in the pilot project and for various inversion methods.

<b>Inversion</b>	<b>Main Results</b>		<b>PDL 150</b>		<b>PDL 300</b>	
	<b><u>pre</u></b>	<b><u>post</u></b>	<b><u>pre</u></b>	<b><u>post</u></b>	<b><u>pre</u></b>	<b><u>post</u></b>
	<u>Information Shares (midpoint)</u>					
HI	0.062	0.321	0.056	0.303	0.062	0.305
LIV	0.032	0.118	0.034	0.112	0.065	0.124
MLIV	0.065	0.331	0.061	0.314	0.067	0.317
	<u>Component Shares</u>					
HI	0.146	0.456	0.147	0.427	0.159	0.427
LIV	0.072	0.164	0.078	0.154	0.110	0.170
MLIV	0.151	0.469	0.153	0.441	0.166	0.443
	<u>Proportion p-value &lt; 5% for Component Shares</u>					
HI	0.432	0.938	0.238	0.738	0.222	0.663
LIV	0.372	0.795	0.179	0.419	0.122	0.198
MLIV	0.437	0.943	0.244	0.754	0.224	0.682



Figure 1. Example of the Implied Prices

The figure plots the implied prices for the LIV, HI and MLIV inversion methods as well as the prices of the underlying security and the call option. The observed prices are for Apple stock and a February ATM call option (strike 135) observed on January 25, 2008. Graphs have been shifted to facilitate presentation.

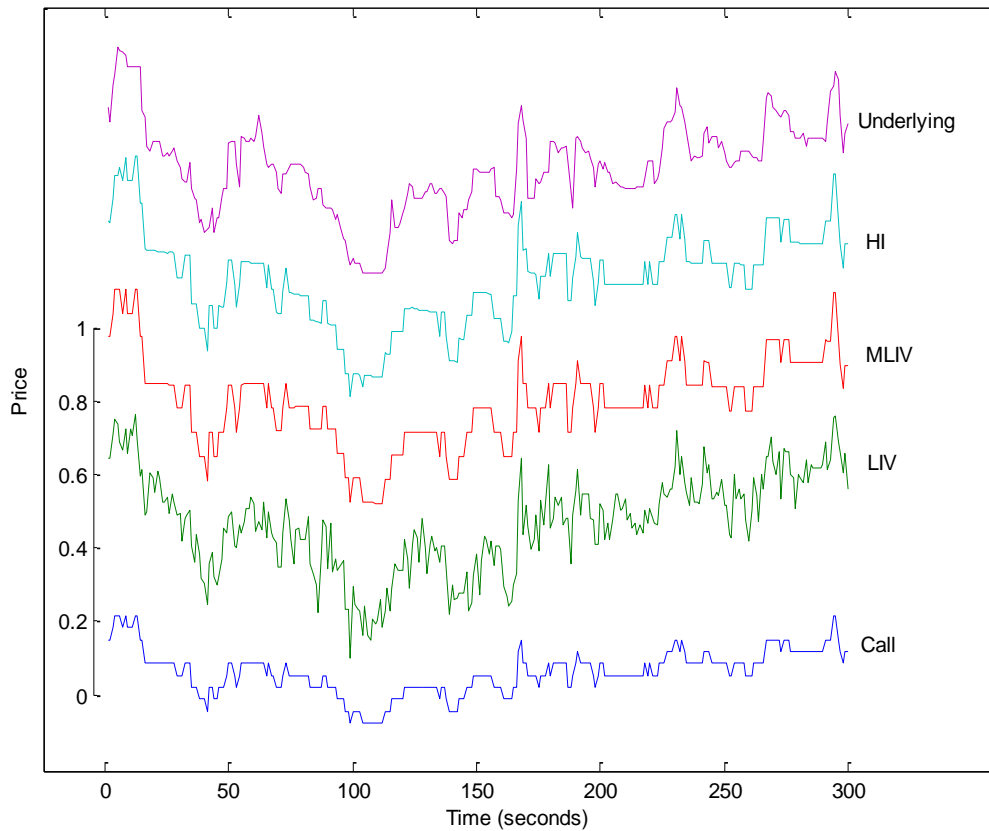
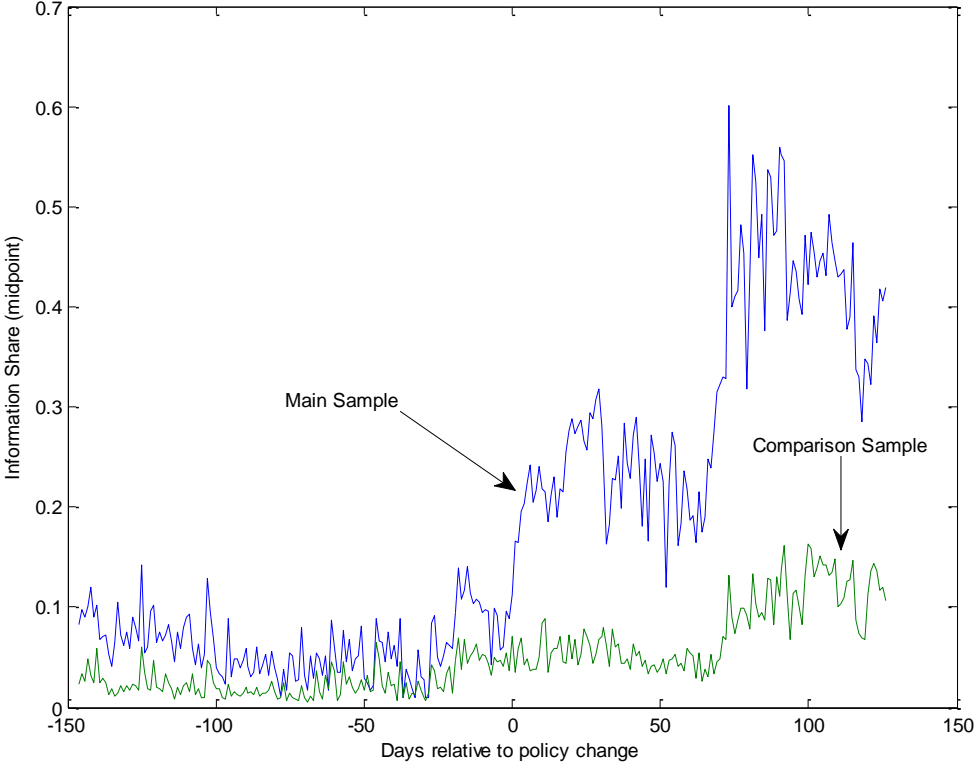


Figure 2: Time Series of Average Midpoint IS



## Appendix

### *Proof of Proposition 1*

Under the error independence assumption we write the structural model (2)-(3) under the form of a matrix equation

$$\begin{bmatrix} \Delta P_t \\ v_t \end{bmatrix} = \begin{bmatrix} v_t + \varepsilon_{S,t} - \varepsilon_{S,t-1} \\ v_t + \varepsilon_{I,t} - \varepsilon_{I,t-1} \\ v_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_t \\ \varepsilon_{S,t} \\ \varepsilon_{I,t} \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{t-1} \\ \varepsilon_{S,t-1} \\ \varepsilon_{I,t-1} \end{bmatrix} \equiv \mathbf{B} \begin{bmatrix} v_t \\ \varepsilon_{S,t} \\ \varepsilon_{I,t} \end{bmatrix} + \theta \begin{bmatrix} v_{t-1} \\ \varepsilon_{S,t-1} \\ \varepsilon_{I,t-1} \end{bmatrix} \quad (\text{A1})$$

Define now

$$\mathbf{B} \begin{bmatrix} \varepsilon_{S,t} \\ \varepsilon_{I,t} \\ v_t \end{bmatrix} \equiv \begin{bmatrix} e_{St} \\ e_{It} \\ e_{vt} \end{bmatrix} = \mathbf{E}_t \Rightarrow \begin{bmatrix} \varepsilon_{S,t} \\ \varepsilon_{I,t} \\ v_t \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} e_{St} \\ e_{It} \\ e_{vt} \end{bmatrix} = \mathbf{B}^{-1} \mathbf{E}_t, \text{ where } \mathbf{B}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}. \quad (\text{A2})$$

Then (A1) can be written under the following Vector Moving Average (VMA) lag polynomial representation, with  $I$  denoting the identity matrix

$$\begin{bmatrix} \Delta P_t \\ v_t \end{bmatrix} = I \mathbf{E}_t + \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{E}_{t-1} = (I + \theta(L)) \mathbf{E}_t. \quad (\text{A3})$$

From (A1)-(A3) the covariance matrix  $\Omega$  of the error vector  $\mathbf{E}_t$  is equal to

$$\Omega = \text{Var} \mathbf{E}_t = \begin{bmatrix} \sigma_S^2 + \sigma_v^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_I^2 + \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 \end{bmatrix}, \quad (\text{A4})$$

From (A3) the identical rows  $\psi$  of the matrix  $\Psi$  in (5) and (6) are equal to  $\psi = [0 \ 0 \ 1]$ . From (5) and (A4) we get  $\psi \Omega \psi' = \sigma_v^2$ , the variance of the random walk.

Since the variance  $\hat{\sigma}^2$  of  $\hat{\varepsilon}_{It}$  forms part of  $\sigma_I^2$ , its effect on the IS bounds for  $I_t$  can be found by exploiting the symmetry of the vector  $\Delta P$  and the matrix  $\Omega$  with respect to the two components of

the vector  $(S_t, I_t)$ . Recall that the IS upper (lower) bound of a price is derived when that price occupies the top (bottom) position in the VMA representation.<sup>41</sup> Symmetry here implies that it suffices to find the impact of the variances  $\sigma_S^2$  and  $\sigma_I^2$  on IS from the Cholesky factorization of (A4), given that  $\Omega$  is not diagonal. Setting  $\Omega = FF'$ , where  $F$  is the lower triangular matrix, we note from the form of  $\psi$  that it suffices to consider the last row of  $F$ , denoted by  $a_{31} \ a_{32} \ a_{33}$ .

Hence, the effect of  $\hat{\sigma}^2$  on the upper bound of IS has the sign of  $\frac{\partial a_{31}^2}{\partial \sigma_S^2}$ ; since  $a_{31}^2 = \frac{\sigma_v^2}{\sigma_S^2 + \sigma_v^2}$  this sign turns out to be negative, QED. The effect of  $\hat{\sigma}^2$  on the lower bound, on the other hand, has the sign of  $\frac{\partial a_{32}^2}{\partial \sigma_I^2}$ . The expression for  $a_{32}$  is complicated and is left as an exercise, but it turns out to be clearly a decreasing function of  $\sigma_I^2$ . This completes the proof for the case of the error independence assumption.

A similar proof also holds for the Hasbrouk-Roll structural model (2)-(8), except for the fact that it involves heavier computations. (A1) now becomes

$$\begin{aligned} \begin{bmatrix} \Delta P_t \\ \Delta E_t \end{bmatrix} &= \begin{bmatrix} u_t + (\lambda + c_S)w_t - c_S w_{t-1} \\ u_t + (\lambda + c_I)w_t - c_I w_{t-1} + \hat{\varepsilon}_{I,t} - \hat{\varepsilon}_{I,t-1} \\ u_t + \lambda w_t \end{bmatrix} \\ &= \begin{bmatrix} 1 & \lambda + c_S & 0 \\ 1 & \lambda + c_I & 1 \\ 1 & \lambda & 0 \end{bmatrix} \begin{bmatrix} u_t \\ w_t \\ \hat{\varepsilon}_{I,t} \end{bmatrix} + \begin{bmatrix} 0 & -c_S & 0 \\ 0 & -c_I & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_{t-1} \\ w_{t-1} \\ \hat{\varepsilon}_{I,t-1} \end{bmatrix} \equiv B \begin{bmatrix} u_t \\ w_t \\ \hat{\varepsilon}_{I,t} \end{bmatrix} + \theta \begin{bmatrix} u_{t-1} \\ w_{t-1} \\ \hat{\varepsilon}_{I,t-1} \end{bmatrix}. \end{aligned} \quad (A1')$$

Similarly, (A2), (A3) and (A4) now become, with  $Var(u_t) = \sigma_u^2$ ,  $Var(w_t) = \sigma_w^2$ ,  $Cov(u_t, w_t) = 0$

$$B \begin{bmatrix} u_t \\ w_t \\ \hat{\varepsilon}_{I,t} \end{bmatrix} \equiv \begin{bmatrix} e_{St} \\ e_{It} \\ e_{qt} \end{bmatrix} = E_t \Rightarrow \begin{bmatrix} u_t \\ w_t \\ \hat{\varepsilon}_{I,t} \end{bmatrix} = B^{-1} \begin{bmatrix} e_{St} \\ e_{It} \\ e_{qt} \end{bmatrix} = B^{-1} E_t, \text{ where } B^{-1} = \begin{bmatrix} -\frac{\lambda}{c_S} & 0 & \frac{(\lambda + c_S)}{c_S} \\ 1 & 0 & -\frac{1}{c_S} \\ -\frac{c_I}{c_S} & 1 & -\frac{c_S - c_I}{c_S} \end{bmatrix} \quad (A2')$$

<sup>41</sup> See Hasbrouk (2002).

$$\begin{bmatrix} \Delta P_t \\ \Delta E_t \end{bmatrix} = I E_t + \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} E_{t-1} = (I + \theta(L)) E_t \quad (\text{A3}')$$

$$\Omega = \text{Var} E_t = \begin{bmatrix} \sigma_u^2 + (\lambda + c_S)^2 \sigma_w^2 & \sigma_u^2 + (\lambda + c_S)(\lambda + c_I) \sigma_w^2 & \sigma_u^2 + \lambda(\lambda + c_S) \sigma_w^2 \\ \sigma_u^2 + (\lambda + c_S)(\lambda + c_I) \sigma_w^2 & \sigma_u^2 + (\lambda + c_I)^2 \sigma_w^2 + \hat{\sigma}^2 & (\lambda + c_I) \sigma_w^2 \\ \sigma_u^2 + \lambda(\lambda + c_S) \sigma_w^2 & \sigma_u^2 + \lambda(\lambda + c_I) \sigma_w^2 & \sigma_u^2 + \lambda^2 \sigma_w^2 \end{bmatrix} \quad (\text{A4}')$$

From (A3') the rows of  $\Psi$  are again  $\psi = [0 \ 0 \ 1]$ , with  $\psi \Omega \psi' = \sigma_u^2 + \lambda^2 \sigma_w^2$ , the random walk variance, as expected.

The main difference between the two structural models is that the vector  $\Delta P$  and the matrix  $\Omega$  are no longer symmetric with respect to the two components of the vector  $(S_t, I_t)$ . As a result, the Cholesky factorization of (A4') as it now stands can only demonstrate the effect of  $\hat{\sigma}^2$  on the IS lower bound, and we need to interchange the rows of the matrix  $\mathbf{B}$  and repeat the estimation in order to find the effect on the IS upper bound. Both factorizations and demonstrations are straightforward but tedious and are left as an exercise.

#### *Proof of Lemma 1*

Recall that the goal is to derive the volatility  $\sigma_t(I_t)$  for the implied price implied by the observed option dynamics. Let  $\sigma_t(df/f)$  denote the instantaneous volatility of the derivative return. By deriving the quadratic variation  $\langle df/f_t, df/f_t \rangle$  for the stochastic differential of the option price (13) and canceling the terms of order higher than  $dt$  we have:

$$\sigma_t^2(df/f) = \left[ S_t f_t^{(S)} \sigma_t^2 + 2\rho S_t f_t^{(S)} \sigma_t f_t^{(\sigma)} \kappa_t + f_t^{(\sigma)} \kappa_t^2 \right] / f_t^2, \quad (\text{A1})$$

which is a quadratic equation in  $\sigma_t$ . The positive root of this equation yields the desired solution:

$$\begin{aligned} \sigma_t(I_t) &= \frac{\sqrt{f_t^{(S)2} \left[ f_t^2 \sigma_t^2(df/f) - 1 - \rho^2 f_t^{(\sigma)} \kappa_t^2 \right]}}{S_t f_t^{(S)2}} - \rho \frac{f_t^{(S)} f_t^{(\sigma)} \kappa_t}{S_t f_t^{(S)2}}, \\ &= \frac{\sqrt{f_t^2 \sigma_t^2(df/f) - 1 - \rho^2 f_t^{(\sigma)} \kappa_t^2}}{S_t |f_t^{(S)}|} - \text{sign } f_t^{(S)} \rho \frac{f_t^{(\sigma)} \kappa_t}{S_t |f_t^{(S)}|} \end{aligned} \quad (\text{A2})$$

where in the first line we preserved the terms whose signs may change upon simplification. Since with our data we may only estimate (A2) under  $\kappa_t = 0$ , which corresponds to univariate diffusion, we gauge the impact of neglecting other terms by binding this impact from above. It is clear from (A2) that stochastic volatility affects  $\sigma_t(I_t)$  the most at the extremes of the admissible values for  $\rho$ . Assuming this quantity to be negative in line with leverage theory and with most empirical studies of stochastic volatility models, we note that an upper bound of the second term of (A2) can be found for  $\rho = -1$ . Approximating the partials in this second term of (A2) by the Black-Scholes (1973) model quantities derived at the ‘appropriate’ implied volatility, we get this term ( $\equiv A$ ) for call options as:

$$A = -\frac{\phi d_1 \sqrt{T} \kappa_t}{\Phi d_1} , \quad (\text{A3})$$

where  $d_1$  denotes the Black-Scholes quantity evaluated at the option implied volatility (IV). Since the ratio of the standard normal density function to the equivalent cumulative probability function is strictly positive and decreasing in its argument, it may be easily shown that  $A$  increases in the  $K/S$  ratio (for calls) and the time to maturity  $T$  while it decreases in IV. By taking the maxima for  $K/S$  and  $T$  in our data, which respectively are 1.1 and 1/3, and the minimum for IV, which is approximately 0.2, we find the approximate extremal value for  $A$  equal to  $-0.7\kappa_t$ . Note that the volatility of the volatility  $\kappa$  is clearly model-dependent. For instance Christoffersen *at al.* (2010) estimate this quantity in the range 0.5-12.5% for the S&P 500 index at different specifications; yielding extremal values of  $A$  between 0.35 and 8.75%. These values overstate the impact of the second term of (A2), not only because  $|\rho| < 1$  but also because our final sample of option contracts contains only the three contracts nearest to the money. For this reason we expect the coefficient of  $\kappa_t$  to be significantly lower in magnitude for our data than its stated extremum of 0.7. Hence we find it plausible that the stochastic volatility will have all but a limited impact on the estimate of the volatility of the implied price for our data.

To evaluate (A2) with the use of observables without the stochastic volatility terms, we note that, since we don’t observe instantaneous quantities, we replace the remaining terms of (A2) by the median quantities for each daily contract-sample. Since the estimate for the partial  $f_t^{(S)}$  from the first differences of the option and underlying would be extremely noisy because of the bid-ask spreads, we use the homogeneity assumption and replace this quantity by  $f_t - f_t^{(X)} X / S_t$ , with  $f_t^{(X)}$  estimated by (14), which factors out the underlying price. Now we are ready to state our approximation for the volatility of underlying price implied by the observed option dynamics:

$$\hat{\sigma} I \cong \tilde{\gamma} \hat{\sigma} df / f , \quad (\text{A4})$$

where for a given contract-day,  $\tilde{\gamma}$  represents the daily median of  $f/|f - f^{(X)}X|$ , with  $f^{(X)}$  estimated as before and we use the square root of the realized variance<sup>42</sup> as the estimate for the option return volatility. Since (A3) is an approximation, we cannot test formally for the differences between this benchmark and the observed volatility of the implied price derived by the inversion methods applied in this study; however, we infer that a ‘good’ inversion method will result in a tight spread around this benchmark quantity.

### *Proof of Lemma 2*

We demonstrate that a powerful benchmark exists for the correlation of the implied price and underlying price,  $\rho_{IS}$ , that is  $\rho_{IS} \geq |\rho_{fS}|$ . In other words, we may use the absolute observed correlation between the option and underlying returns  $|\hat{\rho}_{fS}|$  in statistical tests for the null hypothesis  $\hat{\rho}_{IS} - |\hat{\rho}_{fS}| \geq 0$ . Under a univariate diffusion (including state-dependent volatility), we clearly have  $\rho_{IS} = |\rho_{fS}| = 1$ , which in this case clearly implies that any empirically observed correlation between option and the underlying should hold in its absolute value for the correlation between the implied and underlying price. In the presence of stochastic volatility, by taking the appropriate quadratic and cross variations of the state dynamics (9) and the stochastic differential of the option price (10) and by simplifying we have:

$$\rho_{fS} = \frac{\langle df, dS \rangle}{\sqrt{\langle df, df \rangle} \sqrt{\langle dS, dS \rangle}} = \frac{a + \rho b}{\sqrt{\langle df, df \rangle}} = \frac{a + \rho b}{\sqrt{a^2 + 2\rho ab + b^2}}, \quad (\text{A5})$$

where  $a = S_t f_t^{(S)} \sigma_t$  and  $b = f_t^{(\sigma)} \kappa_t$ .<sup>43</sup> It immediately follows that  $|\rho_{fS}| \leq 1$  since the denominator exceeds  $(a + \rho b)^2$ .

We now prove Lemma 2 by contradiction. Assume  $\rho_{IS} < |\rho_{fS}|$ . By inverting the stochastic differential of the option price (10) for the implied price  $I$ , we have

$$dI_t = \frac{1}{f_t^{(S)}} df + \frac{f_t^{(\sigma)}}{f_t^{(S)}} d\sigma_t + \text{other terms}, \quad (\text{A6})$$

where ‘other terms’ clearly don’t contain either Brownian motion. By taking the appropriate quadratic and cross variation it follows that

---

<sup>42</sup> To estimate the realized variance, the relevant literature recommends using longer time intervals than the one-second frequency because of the microstructure noise present in high-frequency data (see Andersen *at al.* (2002) for a survey). In our estimate we use this one-second frequency since the microstructure noise clearly plays a role in our estimates for the implied price.

<sup>43</sup> Formally, we deal with correlations of return. However, dividing the stochastic differentials in quadratic and cross variations by the appropriate quantities to derive correlation has no effect since those quantities immediately factor out.

$$\rho_{IS} = \frac{\langle dI, dS \rangle}{\sqrt{\langle dI, dI \rangle} \sqrt{\langle dS, dS \rangle}} = \text{sign } f_t^{(S)} \frac{\rho_{fS} - \rho b / \sqrt{\langle df, df \rangle}}{\sqrt{a^2 / \langle df, df \rangle}}, \quad (\text{A7})$$

where  $a$  and  $b$  are as in (A5), in the second equality we separated the terms in the quadratic and cross variations, divided the numerator and denominator by  $\sqrt{\langle df, df \rangle}$  to separate  $\rho_{fS}$  and simplified. Hence, for a call option the relation (A7) implies by assumption

$$\rho_{fS} > \frac{\rho_{fS} - \rho b / \sqrt{\langle df, df \rangle}}{a / \sqrt{\langle df, df \rangle}} = \frac{(a + \rho b) / \sqrt{\langle df, df \rangle} - \rho b / \sqrt{\langle df, df \rangle}}{a / \sqrt{\langle df, df \rangle}} = 1, \quad (\text{A8})$$

where we used  $f_t^{(S)} > 0$  for calls and substituted (A5) in the RHS. Since this last equality contradicts  $|\rho_{fS}| \leq 1$ , it follows that  $\rho_{IS} \geq |\rho_{fS}|$ , QED. The proof for put options follows by extension, QED



## Bibliography

- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2002). Parametric and nonparametric measurement of volatility. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland.
- Anshuman, V.R. and A. Kalay (1998). Market Making with Discrete Prices. *The Review of Financial Studies*, Vol. 11, No. 1, 81-109
- Bacidore, J.M. (1997). The impact of decimalization on market quality: An empirical investigation of the Toronto Stock Exchange. *Journal of Financial Intermediation*, 6, 92-120.
- Bacidore, J. M. and G. Sofianos (2002). Liquidity Provision and Specialist Trading in NYSE-Listed non-U.S. Stocks. *Journal of Financial Economics*, Vol. 63, 133-158
- Bergman, Y. Z., B. D. Grundy, and Z. Wiener (1996). General Properties of Option Prices. *Journal of Finance*, Vol. 51, 1573-1610.
- Berkowitz, J. (2010). On Justifications for the ad hoc Black-Scholes Method of Option Pricing. *Studies in Nonlinear Dynamics & Econometrics*, Vol. 14.
- Bessembinder, H. (2003). Trade execution costs and market quality after decimalization. *Journal of Financial and Quantitative Analysis*, Vol. 38, 747-777.
- Biais, B. and P. Hillion (1994). Insider and Liquidity Trading in Stock and Options Markets. *The Review of Financial Studies*, Vol. 7, No. 4, 743-780
- Bollen, N. and Whaley, R. (1998). Are ‘Teenies’ Better?. *Journal of Portfolio Management*, Vol. 24, 10-24
- Bollen, N. P. B. and Busse, J.A. (2006). Tick size and institutional trading costs: Evidence from mutual funds. *Journal of financial and Quantitative Analysis*, Vol. 41, 915-937.
- Booth, G. G., R. W. So, and Y. Tse (1999), Price discovery in the German equity index derivatives markets, *Journal of Futures Markets*, Vol. 19, 619-643.
- Bowman, A. W. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis*. Oxford University Press, London.
- Capelle-Blancard, G. (2003). Marchés derives et trading de volatilité. *Revue Économique*, No 3, 663-674
- Carr, P. and L. Wu (2009). Variance Risk Premiums. *Review of Financial Studies*. Vol. 22, 1311-1341.

- Chakravarty, S., H. Gulen and S. Mayhew (2004), Informed trading in stock and option markets, *The Journal of Finance*, Vol. 59, 1235-1257.
- Chakravarty, S., Wood, R. and Van Ness, R. (2004). Decimals and liquidity: A study of the NYSE. *Journal of Financial Research*, Vol. 27, 75-94.
- Chen, W.-P., Chou, R. K. and Chung, H. (2009). Decimalization, ETFs and futures pricing efficiency. *Journal of Futures Markets*, 29 (2, 157-178.
- Chen, Y.-L. and Gau, Y.-F. (2009). Tick sizes and relative rates of price discovery in stock, futures and options markets: Evidence from the Taiwan stock exchange. *Journal of Futures Markets*, Vol. 29, No. 1, 74-93.
- Christoffersen, P., Heston, S., and K. Jacobs, 2010, "Option Anomalies and the Pricing Kernel", Working Paper, McGill University.
- Christoffersen P., K. Jacobs, and K. Mimouni (2010). Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices, *Review of Financial Studies* 23, 3141-3189.
- Constantinides, G. M., M. Czerwonko, J. Jackwerth and S. Perrakis (2011). Are Options on Index Futures Profitable for Risk Averse Investors? Empirical evidence. *Journal of Finance*, Vol. 66, No. 4, 1407-1437.
- Dumas, B., J. Fleming, and R. Whaley, 1998. "Implied Volatility Functions: Empirical Tests." *The Journal of Finance*, 53, 2059-2106.
- De Jong, F. (2002). Measures of contributions to price discovery: a comparison, *Journal of Financial Markets*, Vol. 5, 323-327.
- Easley, D., M. O'Hara and P.S. Srinivas (1998). Option Volume and Stock Prices: Evidence on Where Informed Traders Trade. *The Journal of Finance*, Vol. LIII, No. 2
- Eun, C., S. Sabherwal (2003). Cross border listings and price discovery: Evidence from US-listed Canadian stocks. *Journal of Finance*, Vol. 58, 549-575.
- Fleming, J., B. Ostdiek and R. E. Whaley. Trading Costs and the Relative Rates of Price Discovery in Stock, Futures and Option Markets (1996). *Journal of Futures Markets*, 16, 4, 353-387.
- Foucault, T. (1999). Order Flow Composition and Trading Costs in a Dynamic Limit Order Market. *Journal of Financial Markets*, Vol. 2, 99-134.

- Gonzalo, J. (1992). Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics*, Vol. 60, 203-233.
- Gonzalo, J. and C.W.J. Granger (1995). Estimation of common long-memory components in co-integrated systems. *Journal of Business and Economic Statistics*, Vol. 13, 27-36.
- Graham, J.R., R. Michaely and M.R. Roberts (2003). Do price discreteness and transaction costs affect stock returns?: Comparing ex-dividend pricing before and after decimalization. *Journal of Finance*, Vol. 58, 2611-2635.
- Granger, C. W. J. (1980). Testing for Causality: A Personal Viewpoint. *Journal of Economic Dynamics and Control*, Vol. 2, 329-352.
- Harris, L. (1994). Minimum price variations, discrete bid-ask spreads and quotation sizes. *Review of Financial Studies*, Vol. 7, 149-178.
- Harris, F. H. deB., T. H. McInish, G. L. Shoesmith, and R. A. Wood (1995). Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets. *Journal of Financial and Quantitative Analysis*, Vol. 30, 563-579.
- Hasbrouck, J. (1995). One security, many markets: Determining the Contributions to Price Discovery. *Journal of Finance*, Vol. 50, 1175-1199.
- Henker, T. and Martens, M. (2005). Index Futures Arbitrage Before and After the Introduction of Sixteenths on the NYSE. *Journal of Empirical Finance*, Vol. 12, 353-373.
- Heston, S. L., (1993), "A Closed-Form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options." *Review of Financial Studies*, 6, 327-344.
- Johansen, S. (1990). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, Vol.59, 1551– 1580.
- Johansen, S.(1995). Likelihood-based Inference in Cointegrated Vector Autoregressive Models. *Oxford University Press*, Oxford.
- Jones, C. and M. Lipson (2001). Sixteenths: Direct evidence on Institution Trading Costs. *Journal of Financial Economics*, Vol. 59, 253-278.
- Khoury, N., Perrakis, S., Savor, M., (2011), 'Competition, Interlisting and Market Structure in Options Trading', *Journal of Banking and Finance*, Vol. 35, 104-117.
- Kurov, A. (2008). Tick Size Reduction, Execution Costs and Informational Efficiency in the Regular and E-mini Nasdaq-100 Index Futures Markets. *Journal of Futures Markets*, Vol. 28, No. 9, 871-888.

- Normandin M. (2004). Canadian and U.S. Financial Markets: Testing the International Integration Hypothesis Under time-Varying Conditional Volatility. *Canadian Journal of Economic*, Vo. 37, No. 4, 1022-1041
- Pan, J. and A.M. Poteshman (2006). The Information in Option Volume for Future Stock Prices. *The Society for Financial Studies*
- Ronen, T. and Weaver, D.G. (2001). ‘Teenies’ anyone? *Journal of Financial Markets*, Vol. 4, 231-260.
- Sandas, P. (2001). Adverse selection and competitive market making: Empirical evidence from a limit order market. *Review of Financial Studies*, Vol. 14, 705-734.
- Seppi, D.J. (1997). Liquidity Provision with Limit Orers and a Strategic Specialist. *The Review of Financial Studies*, Vol. 10, No. 1, 103-150
- Stephen, J.A. and R.E. Whaley (1990). Intraday Price Change and Trading volume Relations in the Stock and Stock Option Markets. *The journal of Finance*, Vol. XLV, No. 1
- Stock, J. H. and M. W. Watson (1988). Variable Trends in Economic Time Series. *Journal of Economic Perspectives*, Vol. 2, 147–174.
- Yan, B. and E. Zivot (2010). A Structural Analysis of Price Discovery Measures. *Journal of Financial Markets*, Vol. 13, 1-19.