Bachelorthesis Financieel

Jan-Willem Zeijen, 343331

Supervisor: mr. S. Ozturk

May / June 2013

Price Discovery in Stock and Option Market

Abstract

In this paper the issue of price discovery (i.e. the incorporation of new information in the security price) in the stock market on the one hand and the option market on the other hand is analysed. First, call-option values are converted to implied stock prices by using an inverted Black and Scholes (1973) model. Second, the stock-prices are compared to the option implied stock-prices. For this the modified measures of Chakravarty *et al.* (2004), based on the measures of Harris *et al.* (2002) and Hasbrouck (1995), are used. These methods are complemented with the measures from Putnins (2013), to avoid temporary noise effects. Finally, we investigate in time-series determinants of the level of price discovery and we analyse the explicability and predictability of the information shares, following Mizrach and Neely (2008).
Our dataset consists of Apple stock and option bid-/askquotes from March till June 2012.

I Problem description

Nowadays securities of most prominent companies trade not only in the stock exchanges, but also in derivative markets. This research investigates the role of the option market in the incorporation of new information in the price of the asset (i.e. price discovery). If informed traders do trade in the option market, then we will see some adjustment of the stock price due to information available from the option market.

To make a proper analysis on this, we apply the widely used Hasbrouck (1995) method. Although, we modify it slightly to not compare different stock markets with each other, like Hasbrouck (1995) does, but to compare the stock market on the one hand with the option market on the other hand. Second, we inspect the relation between price discovery and variables like spread and volatility. Furthermore we use the adjustment of Putnins (2013) and we analyse the results with regressions proposed by Mizrach and Neely (2008).

II Motivation

Quite some research has been done on price discovery in stock markets. The most prominent research has been done by Hasbrouck (1995). He used common factor models, explained later in this paper, to determine which market is leading in price discovery.
However, the comparison of stock and option markets is an under-analysed research topic (see also Literature).
To prevent from the adverse selection risk, it's relevant for option market makers to know whether the option market is an informative channel. Moreover option market makers may try to hedge via the stock market makers, for who again the informational role of the option market is relevant to know to identify informed trading of the option market makers. Examples from informed trading in the option market can be found in Mayhew, Sarin and Shastri (1995). Pan and Pothesman (2003) give some evidence that trading volume in the option market predicts stock volume.
The given "signals" make us suspect an informational role for the option market, although existing literature is not very unambiguous, sometimes even conflicting, on this role (see Relevance).

To get a more clear insight in the leading / lagging role of the stock / option market, we follow the most prominent paper in this area: it is the one of Chakravarty *et al.* (2004). With their generalized version of Hasbroucks (1995) methodology, we calculate the percentage of price discovery across stock and option markets. However, Chakravarty *et al.* (2004) uses data from 1988 to 1992. Nowadays financial markets have changed in several aspects. Due to, for example, automatic trading and higher internet speed, we now have higher frequency trading possibilities. Therefore in this paper we will review whether the conclusions of Chakravarty *et al.* (2004) are also valid on the Apple stock and option data from 2012.

A problem arising when comparing financial markets, is the temporary noise, coming together with permanent price changes. For this issue, we invest in the elegant solution of Putnins (2013), who uses the method of Harris et al. (2002) in combination with Hasbroucks (1995) method to obtain a robust price discovery measurement.

III Relevance

As stated by Chakravarty *et al.* (2004), there are two relevant contributions of this research. First, some papers conclude that informed trading does not take place in the option markets. These papers try to find which market leads or lags. In some papers the stock market seems to play a leading role when it comes to the impounding of new information in the price, in some papers it doesn't. Conversely, in all relevant papers the authors find no significant leading

role for the option markets. By using Granger lead-lag regressions and similar techniques, some research (e.g. Manaster and Rendleman (1982) and Stephan and Whaley (1990)) tries to detect which market leads or lags when it comes to the impounding of information into prices. On the other hand, more recent research (by Cao *et al.* (2005), Easley *et al.* (1998) and Pan and Poteshman (2006)) reveals options trades could contain information about future stock price movements. This suggests that informed traders *do* trade in option markets. This paper tries to clarify the informational role of option markets in price discovery.
Second, it's interesting to have a look at the different option contracts in terms of leverage, moneyness (spot price divided by strike price) and liquidity, volatility. Which contract has the largest information shares and thus signals best? The goal of this part is, to gain a better understanding of which factors drive the price discovery process for options.

IV Literature

As mentioned, Hasbrouck (1995) suggests an econometric approach based on an implicit unobservable efficient price common to all markets. The information share measures the contribution of the innovations of market j to the innovation in the common efficient price. A lot of studies apply the techniques from Hasbrouck (1995). For example, Baillie (2002) shows that the model of Hasbrouck (1995) and Gonzalo and Granger (1995), who focus on the components of the common factor and the error correction process, yield (nearly) the same results, if residuals are uncorrelated between markets. However, if the residuals are correlated, the upper and lower bounds from Hasbrouck (1995) will diverge further from each other, when correlation is higher. Therefore, the average between upper and lower bound is used as approximation of the markets contribution to the impounding of new information in the efficient price.
Also, in Cuperus *et al.* (2013) Hasbrouck (1995) is used together with the methods of De Jong and Schotman (2010) to analyse the price discovery process in the initial public offering of Facebook stocks on different exchanges. Harris *et al.* (2002) suggests an alternative measure based on Gonzalo and Granger (1995). They use these techniques to analyse the common factor weight which is attributable to three informationally-linked exchanges. The changes can be interpreted as the limits of the changes in the price with respect to the elements of the shock vector, as the time horizon goes to infinity, according to Mizrach and Neely (2008), who compare both standard methods.

As our research focuses on the application of Hasbroucks (1995) techniques on both the stock and the option market, we also mention some relevant papers on this topic. First of all Chakravarty *et al.* (2004) extensively describes how to transform the methods from Hasbrouck (1995) in such a way that these are applicable to our case (see Methodology). Based on five years of stock and options data for 60 firms, they find the option market's contribution to price discovery to be 17% on average. Moreover, they find that option market price discovery is related to trading volume, spread in both markets and stock volatility. Price discovery among different strike prices of an option turns out to be related with leverage, trading volume and spreads. According to the authors, these results suggest an important role for options when it comes to information signalling in price discovery. Another interesting paper is from Chen and Gau (2009), which is an application of Chakravarty *et al.* (2004). They find out that, when analysing price discovery among stock index, index futures and index options in Taiwan with a minimum tick size, the bid-ask spreads of the component stocks of the stock index and the Taiwan Top 50 Tracker Fund get lower, and the contribution of the spot market to price discovery increases.

V Data description

For our research we use Apple data, measured as high frequency tick data, sampled on 1-second frequency. The dataset consist of the stock and options traded on the NASDAQ exchange. This choice is made to avoid noise from using different exchanges and because it follows Chakravarty *et al.* (2004). Moreover, the NASDAQ is by far the biggest index when it comes to Apple data[1] , so that we won't delete much information. Also, Hasbrouck (1995) already concludes only a very small level of price discovery takes place at regional exchanges. Of course, one could extend the research by comparing more stock markets, although taking a too high number of markets will probably lead to vague results and diffused information shares.

From both the stock and option data, we have available the bid- and ask price and the traded volume. The stock bid-/askprices are from the WRDS database, the call-option value are obtained from the Bloomberg Terminal.

We use a sample from 12 March 2012 up to and including 1 June 2012[2] .

VI Data filtering

We filter our data by first deleting all zero bid- and/or ask-quotes and quotes where bid > ask. On the stock bid-/askquotes we run a third filter: we remove quotes when the difference between bid- and ask-quote is bigger than 0.30 USD. Moreover, we delete outliers. We define outliers as bid-/askquotes 4 times the daily standard deviation away from the previous and following bid-/askquotes.

Strictly speaking, applying the third and fourth filter to the implied stock bid-/askquotes would be best. Due to the large compilation time taken by the algorithm to calculate the implied bid-/askquotes, we chose to run the algorithm directly on the midquotes of the call-option prices, to directly obtain midquote stock prices. From empirical research on a small

[1]From the total Apple tradevolume, 60% takes place on the NASDAQ. About 55% of the total number of trades takes place on the NASDAQ.

[2]Due to technical problems data from 16 up to and including 19 April 2012 are not available. This isn't a problem, as the information shares are calculated daily. See Methodology section for details.

piece of the dataset, differences between both methods was very small. This is the main reason for choosing for the faster procedure on midquotes.

Before analyzing the stock and implied option prices, we make sure we only use prices which are non-stationary (i.e. they have a unit root). For this, we use the Augmented Dickey-Fuller (ADF) Test on a 5% significance level, on both the stock and implied option prices (after calculation, see Methodology for calculating implied option prices).
Moreover, we only use stock and implied option prices, which are cointegrated (see Methodology). For this, we apply the Johansen Cointegration Test (1991) on a 5% significance level.

VII Methodology

Hasbrouck (1995) provides a method, based on the co-integration of prices across different stock markets, to calculate each markets information share (IS). The share of market j is defined as the contribution of the variance of market j to the total variance.

Another widely used method is the one from Harris *et al.* (2002), in which the permanent-transitory decomposition framework is adopted. The component share (CS) represents a market's contribution in forming the efficient price innovation.

In our case, we work with a stock and an option market, instead of some stock markets, like Hasbrouck (1995). According to Chakravarty *et al.* (2004), the stock and option market are linked by arbitrage, but there doesn't exist a constant cointegration factor, for the stock and option prices. It is well known, hedge ratios change over time, as a reaction to a changing stock price. On the other hand, we may use an option model to convert option prices into implied stock prices like Manaster and Rendleman (1982) and Stephan and Whaley (1990).

Suppose S_t is the observed value of the stock index. Then, this value can be decomposed in V_t, the implicit, efficient value of the stock index at time t, and an error-term. This error-term is a zero-mean covariance-stationary process:

$$S_t = V_t + e_{s,t}. \tag{1}$$

We can use Black and Scholes (1973) to get an estimate of the call-option price, C_t. Inputs are the risk-free interest rate, maturity of the option, strike and spot price, time till expiration and volatility. For simplicity, only σ, a vector of parameters governing the volatility of the underlying asset. The Black and Scholes formula is denoted by function f:

$$C_t = f(V_t, \sigma). \tag{2}$$

Note that we use the (original) Black and Scholes formula (BS), while dealing with American options. At first sight, this might seem incorrect, because American options have the privilege to be exercised at every point in time till the date of maturity (not just at the date of maturity, like European options), making them having a higher value than European options. However, for American options on non-dividend paying stocks this isn't the case, as there is no reason to exercise the option early, as argued by Geske and Roll (1984)). In fact, Apple pays dividend, but the last cash dividend was paid in 1995 and the the most recent stock-split (stock dividend) took place in 2005 (Apple (2013)). Therefore, using BS on our data sample is justified.

From (2) we can calculate the (needed) implied stock price, I_t:

$$I_t = f_v^{-1}(C_t, \sigma). \tag{3}$$

The problem which arises is all about the unobservable σ. When replacing the constant volatility by the implied time-varying volatility, σ_t, we obtain this volatility by calculating:

$$\hat{\sigma}_t = f_\sigma^{-1}(V_t, C_t). \tag{4}$$

However, this doesn't solve the problem, as V_t is unknown; only S_t is observed. A possible solution is to calculate the implied volatility $\hat{\sigma}_t$ by using lagged values of the index option price and stock index. A drawback of this solution is the fact that the errors over time can be correlated. For that reason, we use a lag of half an hour, as in Chakravarty *et al.* (2004). This should be short enough for the implied volatility to be meaningful, but at the same time long enough to get rid of correlation, although Stephan and Whaley (1990)use volatility of the previous day. We don't choose for this alternative, because volatility may not be constant during the day.

Now we can calculate I_t by using the inverse function of the option pricing formula, with respect to the underlying asset price (k represents the lagging):

$$I_t = f_v^{-1}(C_t, \hat{\sigma}_{t-k}) = f_v^{-1}(C_t, f_\sigma^{-1}(S_{t-k}, C_{t-k})) \tag{5}$$

Now we can introduce the modified method of Hasbrouck (1995). As mentioned, Hasbrouck's (1995) information share represents the relative contribution of a market to price discovery. This can be measured by calculating the contribution of a market to the total variance of the common random-walk component (i.e. the assumed efficient price).

$$p_t = \begin{bmatrix} S_t \\ I_t \end{bmatrix} = \begin{bmatrix} V_t + e_{s,t} \\ V_t + e_{I,t} \end{bmatrix}. \tag{6}$$

In which we assume V_t to follow a random-walk process, with white noise error terms u_t:

$$V_t = V_{t-1} + u_t \tag{7}$$

We then may rewrite this to a Vector Error Correction Model (VECM) of order M:

$$\Delta p_t = A_1 \Delta p_{t-1} + A_2 \Delta p_{t-1} + ... A_M \Delta p_{t-1} + \gamma(z_{t-1} - \mu) + \epsilon_t, \tag{8}$$

with p_t a 2 x 1 vector of prices, A_i a 2 x 2 matrix with the autoregressive coefficients for lag i and error-term $(z_{t-1} - \mu)$ in which $z_{t-1} = p_{1,t-1} - p_{2,t-1}$ and $\mu = E(z_t)$.

We can rewrite the VECM to a VMA (Vector Moving Average) model:

$$\Delta p_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + ..., \tag{9}$$

in which ϵ is a 2 x 1 vector containing the price innovations. This vector has mean zero and variance matrix Ω.

Now, the sum of all moving average coefficients (I is the 2 x 2 identity matrix),

$$\psi(1) = I + \psi_1 + \psi_2 + ..., \tag{10}$$

has identical rows. As ψ reflect the impact of innovations on the permanent price change component (rather than on transient price changes components), we can calculate the total variance of implicit efficient price changes as $\psi\Omega\psi'$. As from Hasbrouck (1995), the information share is the markets proportion in this total variance. If we may assume the Ω matrix to be diagonal (so that the markets innovations are uncorrelated between markets), we can calculate the information share as:

$$S_j = \frac{\psi_j^2 \Omega_{jj}}{\psi\Omega\psi'}. \tag{11}$$

In this expression ψ_j indicates the j^{th} element of ψ and Ω_{jj} represents the $(j,j)^{th}$ element of Ω. As, in practice, markets often *are* correlated, we can only give upper and lower bounds of the informations shares. Correlation appears when the time interval is so long, that price changes and reactions thereon are not measured correctly and are undeserved qualified as contemporaneous. Therefore we should use a 1-second interval. According to Chakravarty *et al.* (2004) we use VAR lags up to 300 seconds.
Of course, there will always rest some correlation. That why Hasbrouck (1995) introduces the Cholesky factorizations ($\Omega = MM'$) to prevent from this correlation. The factorization implies a hierarchy that gives a higher information share for the first price and a lower information share for the last price, in most cases. By performing both the Cholesky factorization and by trying alternative rotations (only 2 in our case), we can get upper and lower bounds.

Because the upper and lower bound of the Hasbrouck (1995) information share estimated from the above model can differ quite a lot, we will use a more in detail defined model for the bivariate case (i.e. comparing the stock to one option). For this purpose, it is easiest to slightly rewrite the VEC model:

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{j=1}^{j=300} A_j \Delta Y_{t-j} + e_t. \tag{12}$$

In this expression from Baillie (2002) *alpha* is the error correction vector and e_t is a zero-mean vector of serially uncorrelated innovations with covariance matrix Ω. This first term of (12) can be interpreted as long-run dynamics between the two timeseries of prices. The second term, indicates the short-term effect caused by market imperfections.

The difference between IS and CS is this specification of the common efficient price. Hasbrouck defines $Y_t = f_t + G_t$, with f_t the common factor and G_t a component with just a temporary effect on Y_t. On the other side, Gonzalo and Granger (1995), on which CS is based, use $f_t = \Gamma Y_t$, in which Γ is the common efficient price coefficient vector. Gonzalo and Granger (1995) show that $\alpha_\perp = (\gamma_1, \gamma_2)'$. Also, it turns out $\beta = (1, -1)'$.
These two results are crucial in rewriting the IS expression for the bivariate case, because $\frac{\psi_1}{\psi_2} = \frac{\gamma_1}{\gamma_2}$. So, we can write:

$$\S_j = \frac{\gamma_j^2 \sigma_j^2}{\gamma_1^2 \sigma_1^2 + \gamma_2^2 \sigma_2^2} \tag{13}$$

However, this equation only holds when there is no significant correlation between the two market's error terms. So, just like in the model described earlier, we use the Cholesky factorization; $\Omega = MM'$. Now one can derive the final expressions for our IS estimates:

$$S_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2} \tag{14}$$

and

$$S_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2} \tag{15}$$

The CS measures are simply computed as: $CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}$ and $CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2}$.

Another way of looking at bivariate markets is presented in Putnins (2013). They conclude that the "conventional" measures like Harris-McInish-Wood (CS) and Hasbrouck (IS) can only draw valid conclusions the first market impounding new information first, when price series have the same level of noise. If this assumption is violated, both measures measure not only the impounding of new information, but also a relative avoidance of noise. It turns out a (relatively) less noise causes a higher information share. This can make the "slowest" market getting the highest information share. Putnins (2013) comes up with a robust metric which combines both measures to get rid of the noise issue in bivariate markets: the informational leadership share (ILS), which is an adaptation, extension and modification of Yan and Zivot

(2010). The model of Yan and Zivot (2010) assumes a fairly simple cointegration model. This makes them observe the IS captures both permanent (a shock leads to the same change in price) and transitory shocks (on the long-run, prices are not effected through the shock), while the CS only captures transitory shocks. On the short-term, affection of price by transitory shocks depends on each price' lag polynomial. So, as Yan and Zivot (2010) conclude, the outcomes may be misleading, as the reaction to temporary shocks differs. To rule out the transitory effects in the measures, the following share is proposed:

$$ILS = \left| \frac{IS_1 CS_2}{IS_2 CS_1} \right| \tag{16}$$

As the above equation shows, the ILS can only be computed in the case of two markets. In our case, we compare options with a strike prices of 500, 505, 510, ..., 695, 700 to the stock-price. We perform this analysis per day to get daily estimates of the CS, IS and ILS measures.

Finally, we compute the Hasbrouck (1995) information shares per day, for both the stock and the option. Following Chakravarty *et al.* (2004) we can investigate in time-series determinants of the level of price discovery by subsampling the information shares on basis of variables as volume and volatility. At this point, we also look at the moneyness of the options, which is defined as S/K, in which S is the stock-price and K is the strike-price of an option.

Also, following Mizrach and Neely (2008), we can analyse the explicability and predictability of the information shares. Therefore we use the average spread during a trading day, the total volume traded on a day and the realized variance (i.e. the annualized daily standard deviation), all for both the stock as the option market as explaining variables. The relevant regression we estimate with OLS is:

$$\begin{aligned} \ln\left(\frac{IS_{j,t}}{1 - IS_{1,t}}\right) &= c + b_1 \cdot \ln\left(\frac{S_{j,t}}{S_{1,t} + S_{2,t}}\right) + b_2 \cdot \ln\left(\frac{N_{j,t}}{N_{1,t} + N_{2,t}}\right) \\ &+ b_3 \cdot \ln\left(\frac{RV_{j,t}}{RV_{1,t} + RV_{2,t}}\right) + b_4 \cdot trend + \epsilon_t \end{aligned} \tag{17}$$

for $j = 1, 2$, which indicates the stock ($j = 1$) or option ($j = 2$) market. S represent the average daily spread, N cointains the daily trading volumes, RV consists of Realized Variances per day and the variable "trend" just represents a simple linear trend. IS contains the estimated information shares per day. This regression is also performed using ILS estimates. HMW regressions are omitted, due to their fluctuating results.

VIII Results

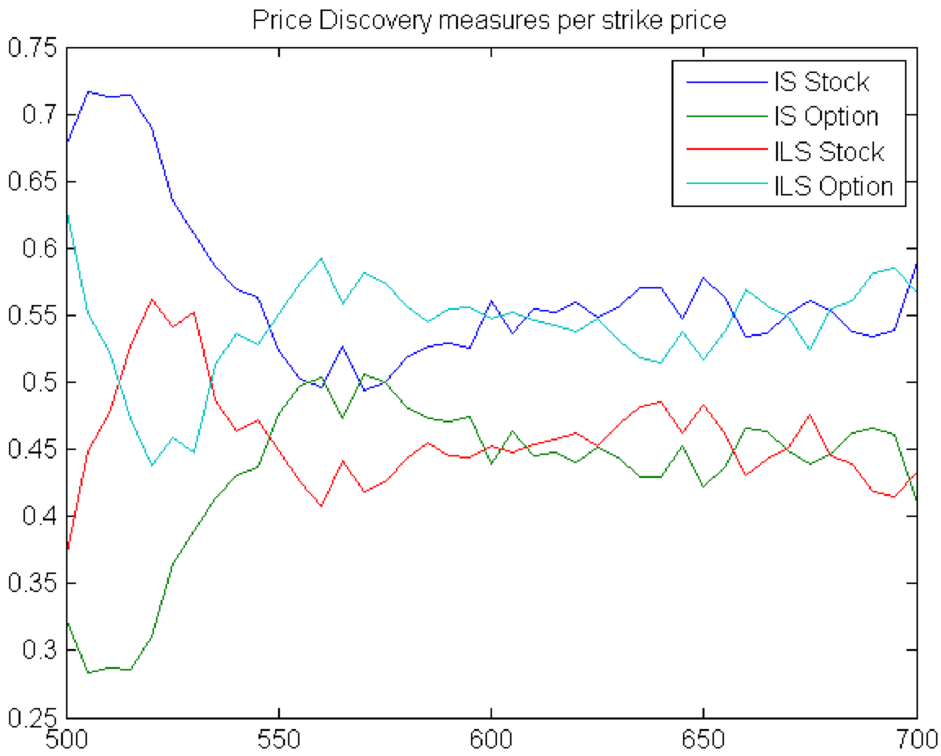

Figure 1: *This figure describes the IS and ILS measures over the different options. The values on the y-axis are the daily estimates of a certain price discovery measure, averaged over time. The x-axis gives the strike price of the option. The moneyness at a strike price of 500 is about 1.17, decreasing to 0.87 at a strike price of 700. The call option with strike price 585 is ATM.*

First of all we convert the call option values to implied stock prices by two times using the Black-Scholes formula; first we compute the implied volatility, thereafter we use the inverse Black-Scholes formula to get an implied stock price. Thereafter, we can compute the I(L)S measures.

For each Apple option (the only difference is the strike price) we take the mean of the daily price discovery measures. So we can easily compare options of different moneyness. These results are displayed in Table 1.

For every option we analyze about 43 trading days in the period 12 March till 1 June 2012.

The HMW measure gives strange results, in the sense that most measures are extremely high or low. The HMW at some options reports acceptable values, but some values are of a far too big magnitude. Therefore, this measure isn't useful.

The Hasbrouck (IS) - we present the mean between upper and lower bound of the IS - mea-

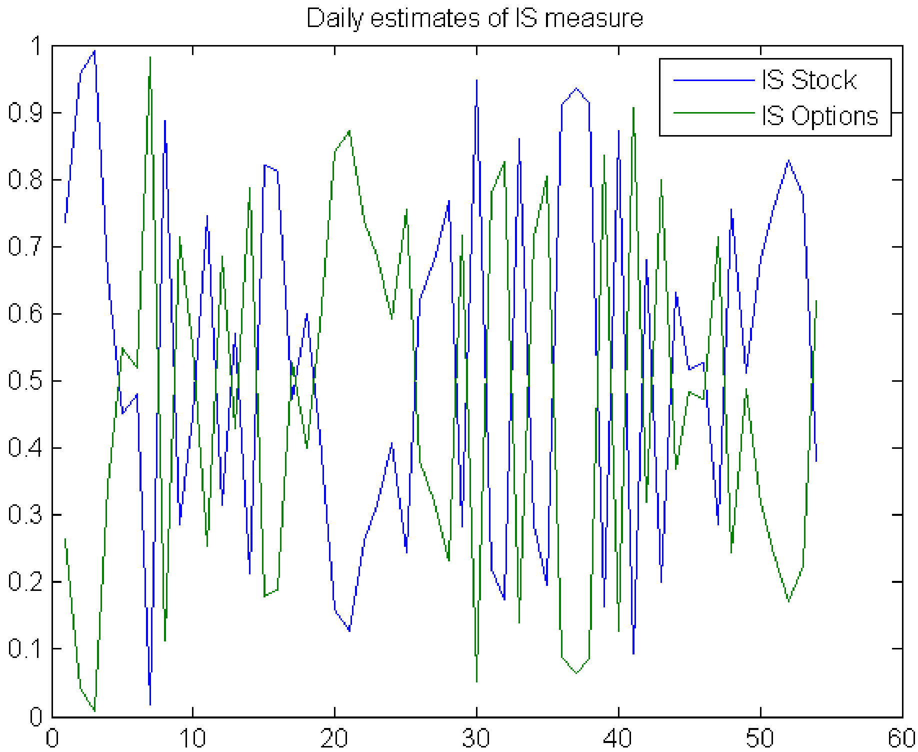

Figure 2: *This figure describes the IS measures over the trading days. The values on the y-axis are the price discovery measures estimated per day, averaged over the options with different strike prices. The x-axis depicts the trading day number.*

sure gives a more stable view on the option price discovery; see also Figure 1. Although, the averaged IS value equals 0.43, which is quite high in comparison with the literature, like Chakravarty *et al.* (2004).

To see whether the somewhat high IS values are a result of the earlier described possible unequal levels of noise between the price series, we perform a calculation of the ILS. The mean of this series is 0.54, which is somewhat higher than values from the existing literature, but only a little higher than the IS measure. This makes it highly convincing, there is relatively less noise in the option market timeseries, causing the IS share of the option market (which seems to be the "slowest" market when compared to the stock market) being relatively high. There is no clear pattern visible in the price discovery measures, although it seems that, according to both measures, the higher strike prices (i.e. the lower moneyness) get a more stable pattern in I(L)S measures. Moreover we have a slightly higher IS / ILS for the options with a relatively low strike price (except from the part around $K = 525$). A possible explanation for this could be the fact that informed traders prefer trading in ITM options, rather than

Table 1: Price discovery measures per option

This table reports the moneyness of each call option, with a given strike price, based on S/K, where S is the average spotprice during the used data sample from 12 March up to and including 1 June 2012. The number of trading days represents the number of days we used to calculate daily estimates of price discovery on the option, compared to the same days of the stock. In the last three columns each column represents one price discovery measure, measures are averaged over the trading days.

Strike price	**Moneyness**	**# trading days**	**HMW**	**IS (mean)**	**ILS**
C500	1.17	9	-0.46	0.32	0.63
C505	1.16	15	-0.10	0.28	0.55
C510	1.15	19	1.38	0.29	0.52
C515	1.14	26	0.26	0.29	0.47
C520	1.13	30	2.54	0.31	0.44
C525	1.12	35	0.72	0.36	0.46
C530	1.11	40	0.15	0.39	0.45
C535	1.10	45	1.88	0.41	0.51
C540	1.09	45	1.38	0.43	0.54
C545	1.08	46	-0.84	0.44	0.53
C550	1.07	48	0.02	0.48	0.55
C555	1.06	49	-0.06	0.50	0.57
C560	1.05	49	0.24	0.50	0.59
C565	1.04	51	-0.11	0.47	0.56
C570	1.03	52	11.65	0.51	0.58
C575	1.02	50	0.30	0.50	0.57
C580	1.01	50	-0.17	0.48	0.56
C585	1.00	50	1.44	0.47	0.54
C590	0.99	50	1.32	0.47	0.55
C595	0.99	52	-0.38	0.47	0.56
C600	0.98	51	36.74	0.44	0.55
C605	0.97	49	-4.77	0.46	0.55
C610	0.96	50	-0.74	0.45	0.55
C615	0.95	50	-0.57	0.45	0.54
C620	0.95	48	-16.71	0.44	0.54
C625	0.94	46	0.93	0.45	0.55
C630	0.93	46	-4.19	0.44	0.53
C635	0.92	42	1.68	0.43	0.52
C640	0.92	43	0.61	0.43	0.51
C645	0.91	43	14.29	0.45	0.54
C650	0.90	42	-0.47	0.42	0.52
C655	0.90	43	-0.01	0.44	0.54
C660	0.89	42	-1.83	0.47	0.57
C665	0.88	41	119.22	0.46	0.56
C670	0.88	41	-0.99	0.45	0.55
C675	0.87	43	1.30	0.44	0.52
C680	0.86	38	-1.16	0.45	0.55
C685	0.86	41	-0.84	0.46	0.56
C690	0.85	41	-0.03	0.47	0.58
C695	0.84	41	0.29	0.46	0.59
C700	0.84	40	-0.04	0.41	0.57

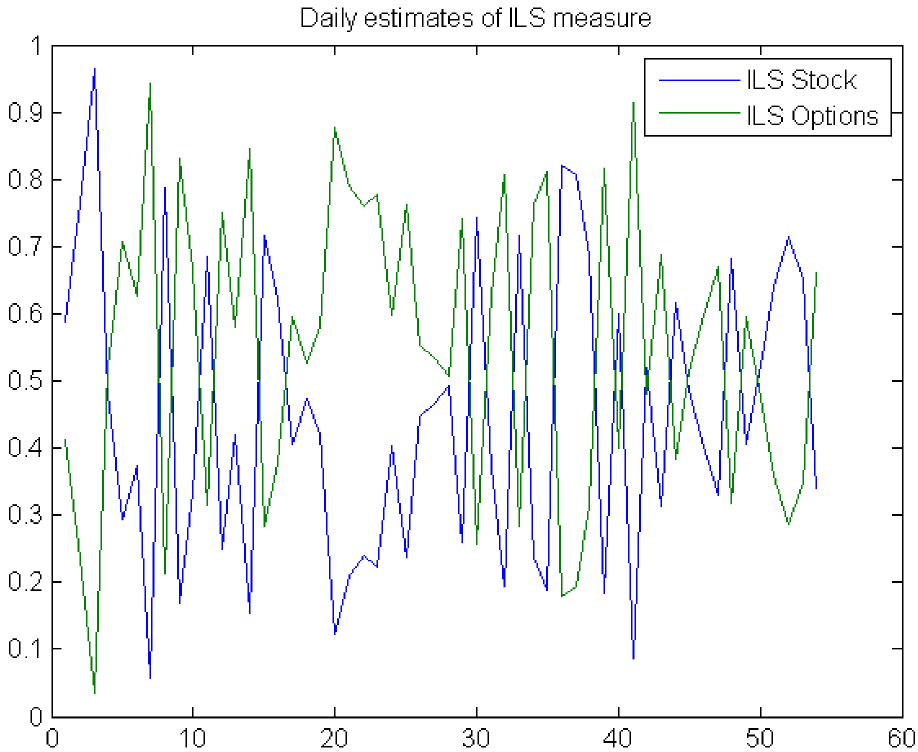

Figure 3: *This figure describes the ILS measures over the trading days. The values on the y-axis are the price discovery measures estimated per day, averaged over the options with different strike prices. The x-axis depicts the trading day number.*

ATM or even OTM options. Therefore, the "signal" they give may be more vague for the higher strike price options, while for options with a lower strike price, there is a clear leading role visible, when compared to the stock market.

A second approach is averaging the estimates of different options per day, so that we can compare different measures over the trading days.
Results are shown in Figure 2 and 3, for respectively the IS and ILS measures.
Again, it seems we capture some noise, due to the fact IS measures both permanent and transitory shocks. While the IS pattern in Figure 2 is heavily fluctuating over time, the graphs in Figure 3 are a little bit more stable.
An overview of both the measures and both stock and option spread is given in Table 2. Also, we publish the daily volatility estimates and the number of trades on both markets.
We perform several regressions on the data from Table 2, but unfortunately don't find any useful relation between I(L)S and the other variables. This could be due to the very fluctuating I(L)S pattern over the days. Perhaps, averaging over more days and using a longer sample

could form a basis for finding this kind of relations.

Table 2: Daily option market Information Shares

This table shows the per day estimates of both the IS and ILS, averaged over the different options. Furthermore, the number of trades in both the stock and option markets are shown. Moreover, we give the daily volatility estimate of the stock market, which is computed as the annualized standard deviation of the returns. Also, the average daily stock spread is presented.

Date	**IS**	**ILS**	**Stock #**	**Opt. #**	**Stock spread**	**Opt. Spread**	**Volatility**
20120312	0.26	0.41	39750	747	0.123	0.223	0.144
20120314	0.04	0.24	141276	4138	0.169	0.536	0.306
20120315	0.01	0.04	121831	4008	0.182	0.737	0.312
20120316	0.34	0.52	72403	1443	0.157	0.411	0.215
20120319	0.55	0.71	90915	2728	0.132	0.438	0.219
20120320	0.52	0.63	82553	1679	0.150	0.385	0.216
20120321	0.98	0.94	62648	1579	0.134	0.372	0.181
20120322	0.11	0.21	61601	1207	0.146	0.374	0.190
20120323	0.71	0.83	47068	790	0.130	0.402	0.145
20120326	0.55	0.67	59636	1191	0.119	0.249	0.140
20120327	0.25	0.32	64639	1677	0.122	0.370	0.159
20120328	0.68	0.75	68000	1914	0.146	0.423	0.180
20120329	0.43	0.58	65326	1331	0.143	0.398	0.171
20120330	0.79	0.85	78620	1652	0.134	0.133	0.173
20120402	0.18	0.28	67592	1745	0.145	0.413	0.154
20120403	0.19	0.38	93562	3559	0.140	0.445	0.188
20120404	0.53	0.60	63636	1779	0.143	0.404	0.165
20120405	0.40	0.53	67873	2201	0.121	0.386	0.142
20120409	0.58	0.58	66035	4972	0.164	0.379	0.180
20120410	0.84	0.88	103960	5901	0.175	0.413	0.228
20120411	0.87	0.79	73535	2560	0.181	0.277	0.203
20120412	0.74	0.76	66433	2152	0.171	0.468	0.191
20120417	0.69	0.78	116974	5525	0.179	0.206	0.299
20120418	0.59	0.60	107326	4783	0.202	0.240	0.277
20120419	0.76	0.76	94979	4741	0.202	0.116	0.289
20120423	0.38	0.55	110322	13896	0.219	0.262	0.412
20120424	0.32	0.54	107247	11902	0.188	0.214	0.307
20120425	0.23	0.51	104995	13187	0.156	0.331	0.236
20120426	0.72	0.74	62396	6090	0.170	0.218	0.177

Continued on next page

Table 2 – *Continued from previous page*

Date	IS	ILS	Stock #	Opt. #	Stock spread	Opt. Spread	Volatility
20120427	0.05	0.26	48475	5786	0.147	0.267	0.151
20120430	0.78	0.63	64701	8546	0.186	0.024	0.217
20120501	0.83	0.81	71536	8139	0.189	0.236	0.230
20120502	0.14	0.28	49349	5475	0.202	0.240	0.217
20120503	0.71	0.76	44262	4101	0.176	0.066	0.188
20120504	0.80	0.81	64808	8847	0.160	0.084	0.198
20120507	0.09	0.18	53192	7701	0.206	0.187	0.231
20120508	0.06	0.19	56396	7791	0.212	0.170	0.243
20120509	0.09	0.32	54418	7861	0.214	0.189	0.231
20120510	0.84	0.82	38132	6753	0.186	0.164	0.188
20120511	0.13	0.40	43557	8063	0.179	0.171	0.186
20120514	0.91	0.91	42329	9882	0.185	0.159	0.202
20120515	0.32	0.48	56575	13253	0.189	0.143	0.233
20120516	0.80	0.69	61843	24843	0.196	0.118	0.258
20120517	0.37	0.38	85056	23104	0.209	0.078	0.280
20120518	0.48	0.51	79747	20821	0.207	0.089	0.291
20120521	0.47	0.60	78984	27423	0.188	0.116	0.255
20120522	0.71	0.67	88759	28575	0.200	0.117	0.270
20120523	0.25	0.32	68543	20273	0.213	0.115	0.306
20120524	0.49	0.60	56369	16947	0.227	0.164	0.271
20120525	0.32	0.47	39068	9446	0.169	0.148	0.204
20120529	0.24	0.36	45172	13680	0.205	0.162	0.216
20120530	0.17	0.29	66566	18546	0.175	0.197	0.223
20120531	0.22	0.35	56933	13619	0.204	0.244	0.224
20120601	0.62	0.66	61710	17718	0.170	0.166	0.211
Average	**0.47**	**0.55**	**71103.91**	**8301.30**	**0.17**	**0.26**	**0.22**

The results of the regression done following Mizrach and Neely (2008) are presented in Table 3. We perform four regressions; for both the IS and ILS measure and for both the stock and option group. We estimate equation (17) with OLS and present the estimated coefficients and corresponding p-values in the table.

There are some conclusions that can be drawn from this output. First, the coefficient for the constant is negative in all cases, which means the other variables overestimate the I(L)S. Second, the spread has a negative coefficient when talking about stocks, as hypothesized by Mizrach and Neely (2008); "a smaller bid-ask spread expedites the tatonnement". The

Table 3: Mizrach and Neely (2008) regression results

Models for both the IS and ILS measures, for the stock as well as for the option. Coefficients estimated by OLS are presented together with their p-values in italics.

	Stock		Option	
	IS	ILS	IS	ILS
Constant	-1.2	-1.06	-9.39	-5.46
	0.421	*0.3179*	*0.048*	*0.107*
Spread	-0.339	-0.159	0.78	1.07
	0.774	*0.8503*	*0.579*	*0.288*
Volume	-153.84	-155.53	-0.9	-0.61
	0.421	*0.256*	*0.105*	*0.153*
RV	-1.13	-0.667	-2.89	-2.22
	0.054	*0.11*	*0.064*	*0.048*
Trend	-0.019	-0.017	0.074	0.051
	0.634	*0.55*	*0.053*	*0.063*
R^2	0.159	0.128	0.218	0.185

spreads have a positive coefficient for the options. The traded volume has a negative effect on the measures; more trading seems to lessen the leading position of the market. This is in contradiction with Mizrach and Neely (2008). The same holds for realized variance; a more volatile markets informational position is relatively lower. This is possibly due to noise trades which (should) diminish the information share. The trend isn't interpretable.
Most variables are not significant on a 5% or 10% level.
The R^2 values are between 0.128 and 0.218. Although we modified the Mizrach and Neely (2008) regression slightly, the magnitude of this numbers is consistent with the magnitude of Mizrach and Neely's (2008) R^2 values.

IX Conclusions

In this research we investigated in the role of price discovery between the stock and option market for Apple, traded on the NASDAQ, in the period March till June 2012.

In his paper Chakravarty et al. (2004) extensively use Hasbrouck's (1995) method to compute information shares for both the stock and option market, by slightly adjusting his techniques.
Putnins (2013) avoids capturing temporary noise effects, by efficiently combining measures of Hasbrouck (1995) and Harris et al. (2002).

We calculated all three measures: Hasbrouck's (1995) information share (IS), Harris' et al. (2002) component share (CS) and Puntin's (2013) informational leadership share (ILS). Over time we see a very fluctuating pattern in both the IS and ILS, between stock and implied option prices. We therefore cannot draw clear conclusions about leading roles over time. Figure 1 shows a reasonably constant IS and ILS. The option I(L)S fluctuates at about 45%, the stock market at 55%. This is not in agreement with the conclusions of Chakravarty et al. (2004), who estimates a much lower (higher) informational role for the option (stock) market.

We had a look at some market related characteristics, but found no useful relationship with the I(L)S measure.

To analyze the predictive power of some market characteristic variables, Mizrach and Neely (2008) invented a regression to predict IS / ILS measures. Due to very fluctuating measures, we can not draw an useful conclusion, although we have acceptable R^2 values and some relations with are understandable.

For further research we would advise to work with a longer time sample (although data availability is less), because daily measures fluctuate heavily. Another possibility is to differ the datafrequency.
The option dataset used by Chakravarty et al. (2004) is big, but a bit outdated. To test the validity of the estimated option market information role in this paper, it would be interesting to not only use a longer time period, but to also have a comparison between multiple companies.

References

[1] Apple Inc., 7 June 2013. *Financial History - Dividend History.* Apple Inc. Investors website.

[2] Baillie, R.T., Booth, G.G., Tse, Y. and Zobotina, T. , 2002. *Price discovery and common factor models.* Journal of Financial Markets 5, pp. 309-322.

[3] Black, F. and Scholes, M., 1973. *The pricing of options and corporate liabilities.* Journal of Political Economy 81, pp. 637-654.

[4] Cao, C., Zhiwu, C. and Griffin, J.M., 2005. *The informational content of option volume prior to takeovers.* Journal of Business, 2005, vol. 78, no. 3.

[5] Chakravarty, S., Gulen, H., Mayhew, S., 2004. *Informed Trading in Stock and Option Markets.* The Journal of Finance 59, pp. 1235-1257.

[6] Chen, Y. and Gau, Y., 2009. *Tick Sizes and Relative Rates of Price Discovery in Stock, Futures, and Options Markets: Evidence from the Taiwan Stock Exchange.* The Journal of Futures Markets 29, no. 1, pp. 74-93.

[7] Cuperus, J., Knecht, W., Volders, K., Zeijen, J., 2013. *Price discovery in fragmented markets.* Werkcollege Financieel 2013.

[8] De Jong, F. and Schotman, P.C., 2010. *Price Discovery in Fragmented Markets.* Journal of Financial Econometrics 8, pp. 1-28.

[9] Easley, D., O'Hare M., Srinivas, P.S., 1998. *Option volume and stock prices: Evidence on where informed traders trade.* Journal of Finance 53, pp. 431-465.

[10] Geske, R. and Roll, R., 1984. *On Valuing American Call Options with the Black-Scholes European Formula.* Journal of Finance 39, pp. 443 - 455.

[11] Gonzalo, J., Granger, C.W.J., 1995. *Estimation of common long-memory components in cointegrated systems.* Journal of Business & Economics Statistics 13, 27 - 36.

[12] Harris, F.H.deB. McInish, T.H. and Wood, R.A., 2002. *Security price adjustment across exchanges: an investigation of common factor components for Dow stocks.* Journal of Financial Markets 5, pp. 277-308.

[13] Hasbrouck, J., 1995. *One Security, Many Markets: Determining the Contributions to Price Discovery.* The Journal of Finance 50, pp. 1175-1199.

[14] Johansen, S., 1991. *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models.* Econometrica 59, no. 6, pp. 1551-1580.

[15] Manaster, S. and Rendleman Jr., R., 1982. *Option prices as predictors of equilibrium stock prices* The Journal of Finance 37, pp. 1043-1057.

[16] Mizrach, B. and Neely, C., 2008. *Information shares in the US Treasury market.* Journal of Banking and Finance, pp. 1221-1233.

[17] Mayhew, S., Sarin, A., Shastri, K., 1995. *The allocation of informed trading across related markets: An analysis of the impact of changes in equity-option margin requirements.* Journal of Finance 55, pp. 1635-1654.

[18] Pan, J. and Poteshman, A.M., 2006. *The information in option volume for stock prices* Review of Financial Studies, vol. 19.3, pp. 871-908.

[19] Putnins, T.J., 2013. *What do price discovery metrics really measure?*. Journal of Empirical Finance, Forthcoming.

[20] Stephan, J. and Whaley, R., 1990. *Intraday price change and trading volume relations in the stock and stock options markets* The Journal of Finance 45,pp. 191-220.

[21] Yan, B. and Zivot, E., 2010. *A structural analysis fo price discovery measures.* Journal of Financial Markets 13, pp. 1-19.