

How Informationally Efficient Are Options Markets? ^{*}

Luis Goncalves-Pinto [†]

Carlo Sala [‡]

April 2, 2022

Abstract

The ability of option-based measures to predict future stock returns is not a sufficient condition for the existence of incremental information in options. If options markets are informationally more efficient than the stock market, then option measures may be used to predict future *actual* stock returns, but there should be weaker or no predictability for future *synthetic* (option-implied) stock returns. We propose to extract the incremental information in option-based measures from their ability to predict the *spread* between actual and synthetic stock returns. We document that existing proxies for informed option trading, such as the option-to-stock-volume ratio, are unable to predict this spread around the release of scheduled and unscheduled firm-specific news. This is evidence inconsistent with the greater informational efficiency of the options markets, and casts doubt on the existence of incremental information in options. The empirical analysis is motivated using a noisy rational expectations model with informed investors who can trade simultaneously in stock and options.

Keywords: Information Markets, Put-Call Parity, Synthetic Returns, Predictability.

JEL Classification: G11, G12, C13.

^{*} For helpful comments and suggestions, we thank Jay Cao, Amit Goyal, David McLean, David Solomon, and seminar participants at the Chinese University of Hong Kong.

[†] University of New South Wales, School of Banking and Finance, Sydney, NSW 2052, Australia; E-mail: l.goncalves-pinto@unsw.edu.au

[‡] Universitat Ramon Llull, ESADE, Department of Financial Management and Control, Avenida de Torreblanca 59, 08172, Barcelona, Spain; E-mail: carlo.sala@esade.edu

How Informationally Efficient Are Options Markets?

Abstract

The ability of option-based measures to predict future stock returns is not a sufficient condition for the existence of incremental information in options. If options markets are informationally more efficient than the stock market, then option measures may be used to predict future *actual* stock returns, but there should be weaker or no predictability for future *synthetic* (option-implied) stock returns. We propose to extract the incremental information in option-based measures from their ability to predict the *spread* between actual and synthetic stock returns. We document that existing proxies for informed option trading, such as the option-to-stock-volume ratio, are unable to predict this spread around the release of scheduled and unscheduled firm-specific news. This is evidence inconsistent with the greater informational efficiency of the options markets, and casts doubt on the existence of incremental information in options. The empirical analysis is motivated using a noisy rational expectations model with informed investors who can trade simultaneously in stock and options.

I Introduction

Can we use option-based measures to forecast future stock returns? Are option markets informationally more efficient than stock markets? Do they provide incremental information that is not already reflected in stock prices? Since the advent of the many existing derivatives markets around the world, these are among the oldest and most debated questions in the literature.

We investigate these questions, both theoretically and empirically, and propose a new method to measure stock price information from options. We argue that the ability of option-based measures to predict *actual* stock returns is *not* a sufficient condition for the existence of incremental information in options. We claim that the appropriate way to measure the information contained in options is not to look at predictability of actual stock returns alone (which is what the majority of existing research does), but to look at predictability of the *difference* between *actual* and *synthetic* (option-implied) stock returns. We find that existing proxies for informed option trading predict both actual and synthetic stock returns to the same extent, around the release of scheduled and unscheduled firm-specific news. This is inconsistent with the idea that options convey *incremental* information that is not already reflected in the underlying stocks.

In the frictionless world of Black and Scholes (1973), markets are complete, and option payoffs can be perfectly replicated by continuously trading the underlying stock and a risk-free bond. In that world, options are redundant securities. However, the real world is plagued with frictions, markets are incomplete, and information is asymmetric. In this world, option payoffs cannot be perfectly replicated, and options are not redundant securities. Thus, informed investors may prefer to trade in options rather than stocks, for three main reasons: (i) the embedded leverage and the truncated payoffs of options (e.g., Back (1993), Biais and Hillion (1994)), (ii) the ability to bet on volatility using options (e.g., Cherian (1993)), and (iii) the short-sale constraints often imposed on stocks (e.g., Lamont and Thaler (2003), Ofek

et al. (2004)).

If informed investors are more likely to trade in options, then one can use option prices and option volume to predict future stock returns. Prior research has shown that option-implied volatilities (Cremers and Weinbaum (2010)), option-implied skews (Xing et al. (2010)), and changes in option-implied volatilities (An et al. (2014)), are all strong predictors of future stock returns. In addition, it has been shown that measures based on option volume can also strongly predict future stock returns, such as the (signed) put-to-call volume ratio (Pan and Poteshman (2006)), and the (unsigned) option-to-stock volume ratio (Roll et al. (2010)). They all argue that the return predictability derived from option metrics is evidence in favour of the hypothesis that the activity in the options market conveys information that is not yet reflected in stock prices.

We argue that the stock return predictability derived from option measures is not a sufficient condition to establish that the options market is informationally more efficient than the stock market, and that there is incremental information in options.

If informed investors do prefer to trade in options, and stock and options markets are not tightly interconnected because of market frictions, then synthetic stock prices should adjust to the fundamental stock values to a larger extent than actual stock prices. Therefore, one should be able to use option signals to predict future *actual* stock returns, but there should be weaker or no predictability regarding future *synthetic* stock returns, depending on the extent of noise trading in options. This second condition has never been examined in prior research, and this project aims to fill this gap. To better understand why this must be the case, Figure 1 illustrates a very stylized example.

[Insert Figure 1 about here]

Panel A of Figure 1 shows the case in which the informed investor prefers to trade in options, and noise trading is absent from the options market. To take advantage of a positive signal received at time $t = 1$, the informed investor purchases call options and sells

put options.¹ This pushes up the synthetic (option-implied) stock price immediately to the fundamental value of the stock. This is because of the assumption that noise trading is absent from the options market. As the new information gets revealed to the stock traders, the actual stock price adjusts to the fundamental value at time $t = 2$. The increase in the synthetic price relative to the actual stock price at time $t = 1$ then signals that there is positive information in the options market not yet reflected in stock prices, and one should buy the actual stock because it is expected to increase in the near future. Moreover, given that the synthetic stock price is already at its efficient level at time $t = 1$, and the fundamental stock value is a random variable, there should be no predictability for synthetic stock returns (i.e., the percent change in synthetic prices).

However, as argued in Grossman and Stiglitz (1980), the situation illustrated in Panel A of Figure 1 can never be an equilibrium, because if the options market is fully efficient, and prices adjust immediately to their efficient levels, this reduces the profitability of informed investors to zero, and they have no incentive to gather any information in the first place. Therefore, there needs to be a certain degree of inefficiency in the options market to encourage informed investors to participate.

Panel B illustrates again the situation of Panel A, but with the presence of noise traders in the options market. Like in Panel A, at time $t = 1$ the informed investor trades on a positive signal by purchasing calls and selling puts, and the synthetic stock price increases towards the fundamental stock value. However, due to the presence of noise traders, prices only adjust partially towards their efficient levels. This means that the informed investor can profit from the trades, and is willing to participate. The increase in the synthetic price of the stock relative to its actual price is a signal that there is positive information in options not yet reflected in stocks, and one should buy the stock as its price is expected to increase in the

¹The case in which the informed investor receives a negative signal at time $t = 1$, and purchases puts and writes calls to exploit such information advantage, can be represented as the mirror image of Figure 1.

near future, i.e., there is predictability in actual stock returns. There is also predictability in synthetic stock returns, because option prices did not adjust fully to their efficient levels at time $t = 1$, and are expected to continue to adjust towards time $t = 2$. However, the predictability in actual stock returns should be stronger than the predictability in synthetic stock returns, for the simple reason that synthetic stock prices have already moved a long way towards the fundamental stock value at time $t = 1$, while the actual stock price still needs to experience a full adjustment towards the fundamental value, like in Panel A.

To distinguish the different degrees of market efficiency across Panels A and B, one could simply examine how much the option signals at time $t = 1$ are able to predict actual compared to synthetic stock returns. In other words, if one *subtracts the synthetic stock return from the actual stock return*, Panel A shows that option signals at time $t = 1$ should predict actual stock returns and the difference between these returns *equally*. However, in Panel B, option signals should be strong predictors of actual stock returns, but *weaker* predictors of the difference in returns. Therefore, the weaker the predictability of the difference in returns, the lower the efficiency of the options market. In the extreme, if option signals fail to predict the difference in returns, this can be interpreted as stock and options markets being equally *inefficient*, and options failing to convey any incremental information relative to the stocks.

Panels A and B assume that the stock and options markets are not fully interconnected. This occurs when there exist market frictions, such as transaction costs, and when options are American-type, which is typically the case for options on individual equities. The two markets are not tightly linked because the put-call parity is not an equality but a set of inequalities (Cox and Rubinstein (1985)). Therefore, the synthetic and actual stock prices can diverge considerably from each other, without leading to arbitrage opportunities. These panels also assume that informed investors choose to trade only in options, which may not always be the case in reality.

Panel C illustrates the case in which the two markets are fully interconnected (no fric-

tions), informed investors can trade in both markets, and there exist noise traders in both markets. This situation can occur when options are European-type, because the put-call parity is an equality, and any deviation from this relation leads to an arbitrage opportunity. Therefore, even if the informed investor prefers to trade in options, the options market maker will arbitrage between the two markets such that the put-call parity will always hold. If the informed investors prefer instead to trade in the stock market, the results would be identical. Therefore, synthetic and actual stock prices are always equal to each other in Panel C, and there is no information in options that is not simultaneously reflected in the stock. This creates a peculiar situation in which the activity observed in the options market at time $t = 1$ can be used to predict future movements in the actual stock price, despite the fact that options do *not* convey any additional information compared to the stock, i.e., options are informationally irrelevant (Chabakauri et al. (2017)). Moreover, by construction, in Panel C the predictability of actual stock returns is exactly the same as that of synthetic stock returns, which results in a zero difference between returns.

Overall, the example illustrated in Figure 1 has two important implications. First, if the options market is informationally more efficient than the stock market, then one should be able to use option signals to predict future actual stock returns, but the predictability of future synthetic stock returns should be weaker (Panel B) or null (Panel A). Second, if the two markets are equally inefficient, and options do not convey any incremental information compared to stocks, then option signals should predict both actual and synthetic stock returns to the same extent, and they should fail to predict the difference (Panel C).

Section III examines this question more formally using a noisy rational expectations model with informed investors who can trade simultaneously in both markets. The framework is borrowed from An et al. (2014), which is an analytic representation of the situation illustrated in Panel C of Figure 1. In the model, there exists an informed investor who can trade simultaneously in a stock, a call option, and a put option, there are noise traders in

both markets, and there is a market maker who arbitrages between the two markets. The options are European-type, which means the put-call parity is a strict equality, and the market maker ensures that this relation is never violated. Therefore, by construction, in the model of An et al. (2014), the synthetic and actual stock prices are always identical. The implications derived from Panel C of Figure 1 also apply to this more formal setting. In particular, the model shows that one can use option signals to predict future actual stock returns when, by construction, options are in fact informationally irrelevant. Therefore, the existence of predictability is not a sufficient condition for the existence of incremental information in options.²

This paper takes these implications to the data and reports surprising findings. It shows that existing empirical proxies for informed option trading, such as the option-to-stock volume (O/S) ratio of Roll et al. (2010), systematically fails to predict the *difference* between actual and synthetic stock returns.³ As discussed above, this evidence is less consistent with the greater informational efficiency of the options market, and is more consistent with the idea that options do not convey incremental information.

The empirical analysis focuses on the periods around the release of scheduled and unscheduled firm-specific news. It uses earnings announcements as scheduled news, and non-earnings 8-K filings as unscheduled news. The period around the release of material corporate news is one in which investors should exhibit stronger incentives to gather private information and trade on this information. The empirical tests show that the O/S ratio on the day

²An et al. (2014) do recognize this serious limitation of their model in their Appendix A (Section A.5). They state that “the predictability of options does not exist after controlling for past stock returns”. However, they argue that in their empirical analysis they do control for past stock returns and the predictability of options does not disappear. However, it could be the case that past stock returns are not a sufficient statistic to capture informed trading in the stock market to explain the predictability of options. Figure 1 (Panel C) proposes an alternative way to adjust the predictability of options. Instead of controlling for past stock returns, one could subtract the synthetic from the actual stock return and test whether option signals predict this difference. If it does, then the options market is informationally more efficient than the stock market, and there is incremental information conveyed by options.

³The empirical analysis focuses on the predictability derived from option volume-based measures, instead of option price-based measures, as the primary explanation for the predictability of the latter group of measures is likely to be price pressure in the stock market, as argued in Goncalves-Pinto et al. (2020).

before an earnings announcement is a strong predictor of actual stock (mid-quote) returns during the announcement window.⁴ However, it is also a strong predictor of synthetic stock returns. In fact, the null hypothesis that actual and synthetic stock returns are equal cannot be rejected in these empirical tests.

Following the implications of the model described above, these results are inconsistent with the idea that the options market is informationally more efficient than the stock market, and that the relative trading activity in options versus stocks conveys incremental information that is not already reflected in stocks.

This paper also examines the predictive ability of O/S on the days of news releases. This test is intended to understand whether option traders are better at processing publicly available information, following Engelberg et al. (2012) and Cremers et al. (2018). The results show that O/S is also a strong predictor of actual and synthetic long-short returns when portfolios are formed on earnings days. The test of the difference between actual and synthetic returns is at times significantly (weakly) different from zero. Overall, the results suggest that option traders may be better at processing information once it is made public, but they do not appear to have an advantage in the anticipation of earnings releases.

The rational model of An et al. (2014) also predicts that in equilibrium the informed investor holding a positive signal is more likely to purchase cheap (out-of-the-money) call options and write expensive (in-the-money) put options. To account for the effect of leverage on the predictive ability of option volume, this paper repeats the analysis described above using the volume-weighted strike-stock price (VWKS) ratio of Bernile et al. (2017). A high (low) value of VWKS indicates that the option volume distribution across strikes is skewed towards the out-of-the-money call (put) options. Therefore, it suggests that in-

⁴ The O/S measure consistently predicts returns negatively. Johnson and So (2012) argue that this could be because informed investors prefer to trade options when holding negative signals. However, the results show that most of the predictive ability of O/S around earnings releases is derived from stocks with low O/S values rather than the stocks with high O/S values.

formed investors holding a positive (negative) signal are looking to leverage up their private information by purchasing out-of-the-money calls (puts).

This measure is also a significant predictor of both actual and synthetic stock returns around earnings announcements. However, the predictability using VWKS seems to be stronger than that of O/S right after the announcement, and weaker than that of O/S a few days after the announcement. This is because the VWKS measure is a function of the underlying stock price, and is therefore more sensitive to stock price pressure, as argued in Goncalves-Pinto et al. (2020).⁵ The null hypothesis that actual and synthetic stock returns are equal cannot be rejected when using the VWKS as predictor, which is inconsistent with the greater informational efficiency of the options market.

Panel C of Figure 1 suggests that these results could be driven by the tight interconnect-
edness between stock and options markets, which does not allow synthetic prices to deviate from actual stock prices significantly. However, the option contracts used in the empirical analysis are generally of the American-type, and bid-ask spreads in the options market are large. This results in a range between the upper and the lower non-arbitrage bounds of Cox and Rubinstein (1985) that is between 1.37% and 2.43% of their mid-point. This means that, on the portfolio formation days, the mid-point between the actual bid and ask stock prices (which are used to compute mid-quote returns) can vary significantly without leading to arbitrage opportunities. Therefore, the linkage between the two markets is not tight, and cannot explain the results reported above.

Lastly, this paper shows that, the results presented above for the period around earnings announcements, are qualitatively similar when focusing on the period around the disclosure of (non-earnings) 8-K filings.

The remainder of this paper is organized as follows. Section II reviews the related liter-

⁵ In a separate test, it is shown that, the predictive ability of O/S does not appear to be sensitive to stock price pressure, while VWKS is very sensitive. This is expected given that this measure is a function of actual stock prices.

ature. Section III describes a noisy rational expectations model with an informed investor who can trade simultaneously in a stock, a call option, and a put option. Section IV provides an empirical analysis of some of the implications of the model. Section V concludes the paper.

II Related Literature

This paper is related to prior research that argues that there is no significant price discovery in the options market. It focuses on option volume-based measures, and on the periods around the release of scheduled and unscheduled firm-specific news. Below is a short description of some of the most closely related existing references. This list is not intended to be exhaustive. It is important to highlight that none of the related references mentioned below uses the predictability of synthetic stock returns to assess the degree of inefficiency of the options market, which is the main contribution of this paper.

There are a number of prior studies that also argue that options do not convey significant information about stock returns. For instance, Stephan and Whaley (1990) and Chan et al. (1993) find no evidence that price changes in the options market lead price changes in the stock market. Vijh (1990) finds that large option trades have relatively small price effects, which is inconsistent with such trades being related to information. Chan et al. (2002) study the cross-market lead-lag effects and show that stock volume has predictive ability for both stock and option prices, but option volume has no incremental predictive ability. Muravyev et al. (2013) show that, when the stock and options markets disagree about the stock value (i.e., a deviation from the put-call parity occurs), it is the option quote that adjusts to correct the disagreement, and the stock quote does not adjust. More recently, Collin-Dufresne et al. (2017) focus their analysis on stock and option trades by Schedule 13D filers and find that volatility information is more likely to be reflected in options, while price information is more

likely to be reflected in stocks.

The studies mentioned above are all empirical in nature. The theoretical research on this topic is more incipient. For instance, Chabakauri et al. (2017) show that adding options to an economy with a single stock (like Grossman and Stiglitz (1980)) does not help reveal any additional information about the distribution of payoffs of the stock, and options are therefore informationally irrelevant. The studies by An et al. (2014) (described in more detail in the next section) and Easley et al. (1998) both fall into this category. Although the latter study is often times cited as predicting a preference by informed traders to trade in options first, their model does not make such a prediction, as any information in the history of stock and option trades is simultaneously reflected in both markets.

More generally, this paper is related to the literature that argues against the superior informational efficiency of the options market. For instance, Cao et al. (2018) show that option returns are predictable using a variety of underlying stock characteristics and firm fundamentals, such as idiosyncratic volatility, past stock returns, profitability, cash holding, new share issuance, and dispersion of analyst forecasts. They also conclude that such predictability patterns have important implications for the efficiency of the options market.

It is common to use material corporate events (and specially earnings announcements) as a way to validate the informed trading proxies. For instance, Amin and Lee (1997) show that option trading predicts stock returns around quarterly earnings announcements, and conclude that option traders participate in price discovery and in the dissemination of earnings news. Roll et al. (2010) show that O/S is higher around earnings announcements. Hu (2014) finds that option order imbalances predict abnormal stock returns five days before earnings announcements (hard to reconcile with the previous two studies). More recently, Cremers et al. (2018) examine the predictive ability of signed option volume around both scheduled and unscheduled corporate news. They find that purchases of options predict returns on news days and ahead of unscheduled events, but not before scheduled events, and

sales of options are informative only ahead of scheduled news releases.

III Theoretical Framework

This section builds on An et al. (2014). They propose a model in which informed investors are allowed to trade simultaneously in stocks and (European-type) options. The extent to which they trade depends on the amount of noise trading present in the two markets. The prices of stocks and options are determined by a market maker who can arbitrage between the two markets. Like in a typical noisy rational expectations model, the prices of stocks and options change via the trading of the informed, but they are not fully revealing, otherwise the informed investors would not profit from gathering information. This creates predictability from option prices and volume to future stock returns. However, by construction, the options in this model do not convey any additional information that is not already reflected in stocks. Therefore, the use of the predictability from option prices and volume to future stock returns as a way to measure the informational advantage of options, is a procedure that needs to be redefined.

A Economy

The firm is born at date $t = 0$, investors trade the stock and options at date $t = 1$, and the firm's cash flows F are realized at time $t = 2$. The distribution of F is binomial, where $F = F_H$ with probability 0.5, and $F = F_L$ with probability 0.5. The prices of the stock at times $t = 0$ and $t = 1$ are denoted as S_0 and S_1 , respectively.

There exist call and put options written on the stock. Their strike price is K , where $F_L < K < F_H$, and the options mature at time $t = 2$. The prices of the call (put) option at times $t = 0$ and $t = 1$ are denoted as C_0 and C_1 (P_0 and P_1), respectively. The payoffs of the call and the put at time $t = 2$ are $C_2 = \max(F - K, 0)$ and $P_2 = \max(K - F, 0)$,

respectively.

There exist informed traders, noise traders, and a market maker, all with CARA utility, and with risk aversion γ . Informed traders receive a signal θ just before date $t = 1$. This signal takes the value $\theta = 1$ with probability of 0.5, and the value $\theta = 0$ with probability of 0.5. If $\theta = 1$, the cash flow of the firm is equal to F_H with probability ω , and it is equal to F_L with the probability $1 - \omega$. If $\theta = 0$, the cash flow of the firm is equal to F_H with probability $1 - \omega$, and it is equal to F_L with probability ω . The parameter ω captures the quality of the signal θ , where $\frac{1}{2} < \omega < 1$. Therefore, the probability that the firm cash flow is F_H , conditional on the signal θ , is given by $p(\theta) = \omega\theta + (1 - \omega)(1 - \theta)$. The conditional probability of a firm cash flow of F_L is therefore $1 - p(\theta)$.

Informed traders can trade both the stock and the options. Their demand for stock is denoted as q_I , and their demands for call and put options are denoted as d_I and u_I , respectively. The corresponding demands for stock and options for the market maker are denoted as q_D , d_D , and u_D , respectively. There exist noise traders in both the stock market and the options market. Their demand for stock is denoted as $z \sim N(0, \sigma_z^2)$, and their demands for call and put options are denoted as $\nu_c \sim N(0, \sigma_c^2)$ and $\nu_p \sim N(0, \sigma_p^2)$, respectively. Noise traders cannot trade across stock and options markets, which means that z is independent of ν_c and ν_p . However, their demands for calls and puts can be correlated. The market-clearing conditions for the stock and options markets are as follows:

$$q_I + q_D + z = 1 \tag{1}$$

$$d_I + d_D + \nu_c = 0 \tag{2}$$

$$u_I + u_D + \nu_p = 0 \tag{3}$$

which means that the options are in zero net supply and there is one share of stock.

B Equilibrium

If the informed investor receives no signal, and there are no demand shocks in either market at time $t = 0$, the informed investor and the market maker are identical, they buy half of the stock each at price S_0 , and there is no trading in options.

If the informed investor receives a signal just prior to time $t = 1$, her objective is to maximize CARA utility from terminal wealth:

$$\max_{q_I, d_I, u_I} E \left[-\frac{1}{\gamma} \exp(-\gamma W_I) \mid \theta \right] \quad (4)$$

subject to

$$q_I S_1 + d_I C_1 + u_I P_1 = \frac{1}{2} S_1 \quad (5)$$

where $W_I = (F - S_1)q_I + (C_2 - C_1)d_I + (P_2 - P_1)u_I$, and γ is the absolute risk aversion of the informed investor. Taking the first order conditions with respect to q_I , d_I , and u_I , gives the following relation:

$$q_I(F_H - F_L) + d_I(F_H - K) - u_I(K - F_L) = -\frac{1}{\gamma} \log \left(\frac{1 - p(\theta)}{p(\theta)} \frac{S_1 - F_L}{F_H - S_1} \right) \quad (6)$$

and the prices of the call and the put at time $t = 1$ are given by:

$$C_1 = \frac{F_H - K}{F_H - F_L} (S_1 - F_L) \quad (7)$$

$$P_1 = \frac{K - F_L}{F_H - F_L} (F_H - S_1) \quad (8)$$

Both the call and the put prices are linear with respect to the stock price. This is because of the assumption of a binomial distribution for the firm cash flows.

The marker maker does not observe the signal θ and is assumed to have unlimited wealth,

which means that she has no budget constraint. The first order conditions for the market maker are as follows:

$$q_D(F_H - F_L) + d_D(F_H - K) - u_D(K - F_L) = -\frac{1}{\gamma} \log \left(\frac{S_1 - F_L}{F_H - S_1} \right) \quad (9)$$

The price of the stock is derived from the sum of (6) and (9), which results in:

$$S_1(\theta, z, \nu_c, \nu_p) = \frac{G(\theta, z, \nu_c, \nu_p)}{1 + G(\theta, z, \nu_c, \nu_p)} F_H + \frac{1}{1 + G(\theta, z, \nu_c, \nu_p)} F_L \quad (10)$$

where the function $G(\theta, z, \nu_c, \nu_p)$ is given by:

$$G(\theta, z, \nu_c, \nu_p) = \sqrt{\frac{p(\theta)}{1 - p(\theta)}} \exp \left(-\frac{1}{\gamma} ((1 - z)(F_H - F_L) - \nu_c(F_H - K) + \nu_p(K - F_L)) \right)$$

The equilibrium stock price (10) is essentially a weighted-average of the high and low firm cash flows to be realized at time $t = 2$, where the ratio $\frac{G(.)}{1+G(.)}$ represents the weight allocated to the high cash flow F_H , the remaining being allocated to the low cash flow F_L . The numerical example below shows that $G(.)$ is typically lower than 1, which makes S_1 skewed towards F_L .

C Numerical Example

Following the calibration used in An et al. (2014), the options are taken to be approximately at the money, i.e., $K = E(F)$, and the baseline parameter values used in this numerical example are as follows: $F_H = 103$, $F_L = 97$, $K = 100$, $\omega = 0.7$, and $\gamma = 1.5$.

Informed investors trade the stock and the two options, and the extent to which they trade depends on the amount of noise trading in the two markets. As they trade to take advantage of their private information, the prices of the stock and the options change at time $t = 1$. Figure 2 plots the prices of the stock, call, and put options, as functions of

noise trading demands, conditional on a positive signal $\theta = 1$ arriving just before time $t = 1$. Panel A plots the price of the stock as a solid line for stock demand shocks z , while keeping the call and put demand shocks at $\nu_c = 0$ and $\nu_p = 0$. The dash-dotted line represents the stock price as a function of call demand shocks ν_c , while holding the stock and put demand shocks at $z = 0$ and $\nu_p = 0$. The dotted line represents the stock price as a function of put demands shocks ν_p , holding the stock and call demand shocks at $z = 0$ and $\nu_c = 0$. Panels B and C repeat this exercise but plot instead the prices of the call and the put options, respectively.

[Insert Figure 2 about here]

Compared to Figure A1 in An et al. (2014), this figure includes the effect of put demand shocks on stock, call, and put prices. As expected, the higher the noise trading demand for the put option, the higher the put price, but the lower the prices of the stock and the call option.

In this example, the put option is significantly more expensive than the call option. The put prices range between 2.88 and 2.98, while the call prices range between 0.02 and 0.12. This is because the equilibrium stock price S_1 in Panel A is very close to the low cash flow F_L . Therefore, at time $t = 1$, the call option is out-the-money, and the put option is in-the-money, given the strike price $K = 100$.

Comparative Statics

Figure 2 shows that the curves intersect at the point in which all the noise trading shocks are equal to zero (i.e., $z = 0$, $\nu_c = 0$, and $\nu_p = 0$). This section examines the sensitivity of that intersection point to changes in some parameter values.

Figure 3 shows the sensitivity of the stock price (Panel A), the call price (Panel B), and the put price (Panel C), to changes in the probability of the high cash flow (ω), given a

positive signal ($\theta = 1$). The parameter ω can also be interpreted as the quality of the signal. Its baseline value is $\omega = 0.7$, and Figure 3 takes ω to vary in the interval $]0, 1[$. As explained in Section III.A, for ω to represent the quality of the signal θ it needs to be bounded (i.e., $\frac{1}{2} < \omega < 1$). Therefore, the results in Figure 3 for $\omega \leq \frac{1}{2}$ can be ignored.

[Insert Figure 3 about here]

Not surprisingly, as the quality of the positive signal ($\theta = 1$) improves (i.e., ω increases), the equilibrium prices of the stock and the call increase, and the put price decreases. The sensitivity of the call price is the strongest. Its price increases by 6 times from 0.05 to 0.3 as the quality of the signal ω changes from its baseline value of 0.70 to 0.99. The price of the put option decreases by 8.5% from 2.95 to 2.70 and the price of the stock increases by only 0.5% from 97.1 to 97.6, for the same change in ω . The reason for this differential sensitivity is that the equilibrium stock price is very close to the low cash flow value of $F_L = 97$. Given the strike price of the options of $K = 100$, the call option is out-of-the-money and the put option is in-the-money. The delta of the call is positive and its gamma increases as the stock price approaches the strike price.

The results in Figure 3 suggest that an informed investor with a positive signal about the stock can profit the most when purchasing the out-of-the-money call option and writing the in-the-money put option. Moreover, these results suggest that there is money to be made from collecting information to improve the quality of the private signal.

Figure 4 examines the relation between equilibrium prices and changes in the risk tolerance of the informed investor (γ), keeping everything else in the baseline model unchanged.

[Insert Figure 4 about here]

The results show that, as the informed investor's risk aversion increases, the prices of the stock and the call decrease, and the price of the put increases. The analytic solution for the equilibrium stock price (10) shows that the value of the function $G(\theta, z, \nu_c, \nu_p)$, which

determines the weight on the high cash flow F_H , decreases as γ increases. Therefore, as γ increases, the stock price converges to the low cash flow $F_L = 97$, the call price converges to zero, and the put price converges to $K - S_1 = 3$.

Lastly, Figure 5 examines the sensitivity of prices to changes in the volatility of the cash flows at the terminal date $t = 2$. The cash flow volatility can be captured by the distance between the high and the low firm cash flows, i.e., $F_H - F_L$. The variation in cash flows is kept symmetric in relation to the strike price $K = 100$, which means that $F_H = K + a$ and $F_L = K - a$. In the baseline model, $F_H - F_L = 103 - 97 = 6$, which corresponds to the case $a = 3$. Figure 5 plots the results for values of a between 0.05 and 5.05.

[Insert Figure 5 about here]

In the baseline case (i.e., $a = 3$), the equilibrium stock price, when noise trading is absent (i.e., $z = 0$, $\nu_c = 0$, and $\nu_p = 0$), is $S_1 = 97.10$ (see Figure 2, Panel A). This is a value close to the low cash flow of $F_L = 97$. The results in Figure 5 (Panel A) show that, as the dispersion in cash flows increases around the strike price (i.e., a increases), the equilibrium stock price decreases because it continues to be close to the low cash flow. For instance, for $a = 2$, $F_L = 98$ and $S_1 = 98.28$, and for $a = 5$, $F_L = 95$ and $S_1 = 95.01$.

For lower values of the cash flows dispersion (a), the equilibrium stock price S_1 converges to the strike price $K = 100$. The gamma of the call option is highest the closer the option is at-the-money. This explains the hump-shaped results in Panel B of Figure 5. It shows that the call price increases from 0.029 to a maximum value of 0.255 when a increases from 0.05 to 0.90. For high values of a , the price of the call decreases to zero. This is because, for high values of a , the call option goes deep out-of-the-money.

Panel C shows the effect of changes in cash flow volatility of put prices. It shows a monotonic increase in this price as the dispersion in cash flows increases. This is because, for a given strike price $K = 100$, higher a leads to lower S_1 , and the put option goes deep in-the-money.

D Predicting Stock Returns with Option Measures

This section discusses the predictability that can be derived from the model described in III.A and III.B. It argues that such predictability is not a sufficient condition for options to carry any incremental information, and proposes an extension that would allow options to convey information that is not already reflected in stocks.

D.1 Single Stock, European Options, and Informational Irrelevance

Chabakauri et al. (2017) show that adding options to an economy with a single stock does not help reveal any additional information about the distribution of payoffs of the stock, and options are therefore informationally irrelevant. The model of An et al. (2014) falls into this category, as well as the highly cited model of Easley et al. (1998). In both models, any information in the history of stock and option trades is simultaneously reflected in both markets.

In the model described in Section III.B, the informed investor can trade a single stock and European-type call and put options written on that stock. The model-implied synthetic stock price, which is denoted as S_1^* , is given by the put-call parity for European-type options:

$$S_1^* = C_1 - P_1 + K \tag{11}$$

and by plugging (7) and (8) into (11), the synthetic stock price is always equal to the actual stock price, i.e. $S_1^* = S_1$. This is true not only at time $t = 1$, but also at time $t = 2$. This is because the market maker arbitrages between the stock and the options markets so that the put-call parity always holds. As a result, and by construction, in this model the actual and synthetic stock returns are always identical. This means that there is no information in option prices that is not also reflected in the stock price.

In the presence of noise trading in stock and options, prices do not adjust to their fully

revealing levels, which leads to predictability from time $t = 1$ to $t = 2$. This means that, the existence of predictability from options to future actual stock returns is not a sufficient condition for the existence of additional information in options that is not also reflected in stocks. An et al. (2014) do recognize this serious limitation of their model in their Appendix A (Section A.5). They state that “the predictability of options does not exist after controlling for past stock returns”. However, they argue that in their empirical analysis they do control for past stock returns and the predictability of options does not disappear. It could be the case that past stock returns are not a sufficient statistic to capture the information reflected in the stock market. Figure 1 (Panel C) proposes an alternative way to adjust the predictability of options. Instead of controlling for past stock returns, one could subtract the synthetic from the actual stock return and test whether option signals predict this difference. If it does not, then it is hard to argue that the options market is informationally more efficient than the stock market, and that there is incremental information in options.

D.2 American Options and Transaction Costs

There can only be incremental information in options if one allows for a looser linkage between the stock and the options markets. In the model described above, the prices of stock and options are very tightly linked by the put-call parity. This is because options are European-type and transaction costs are absent. This is the case illustrated in Panel C of Figure 1.

The model could potentially be extended to take into account the existence of bid-ask spreads, dividends, non-zero interest rates, and American-type options. In this extended model, the put-call parity relation would become a pair of inequalities (Cox and Rubinstein (1985)).⁶ This would lead to a lower and an upper bound on the stock bid and ask prices, and as long as they remain within these bounds there will be no arbitrage opportunities.

⁶ Such an extension is non-trivial and is left for future research.

The lower bound on the stock ask price is given by:

$$S_1^{ask} \geq S^L \equiv C_1^{bid} + Ke^{-rT} - P_1^{ask} \quad (12)$$

which precludes arbitrage from going long on the actual stock and short on the synthetic stock. If this condition does not hold, i.e., $S_1^{ask} < C_1^{bid} + Ke^{-rT} - P_1^{ask}$, then the marker maker (i.e., the arbitrageur) buys the put and the stock, and writes a call and sells a risk-free bond with face value K . This generates a risk-less profit. Indeed, if the buyer of the call exercises early to capture a large dividend (Div): the arbitrageur loses the stock, and the value of her portfolio is $P + S_t - S_t + K - Ke^{-r(T-t)} = P + K - Ke^{-r(T-t)} > 0$. If the dividend (Div) is not large enough to trigger early exercise, the arbitrageur collects the dividend and holds the options until maturity: $P_T + S_T + De^{r(T-t)} - C_T - K = De^{r(T-t)} > 0$.

The upper bound on the stock bid price is given by:

$$S_1^{bid} \leq S^U \equiv C_1^{ask} + K + PV(\text{Div}) - P_1^{bid} \quad (13)$$

which precludes arbitrage from going short the actual stock and going long the synthetic stock. If this condition does not hold, i.e., $S_1^{bid} > C_1^{ask} + K + PV(\text{Div}) - P_1^{bid}$, then the arbitrageur buys a risk-free bond with a face value in the amount of Div, buys a risk-free bond priced at K , writes a put, and shorts the stock. This leads to a risk-less profit. Because she is short the stock, she is responsible for paying the dividends and $PV(\text{Div})$ allows her to meet this obligation. If the buyer of the put exercises early, the value of the arbitrageur's portfolio is $C + Ke^{rt} + S_t - K - S_t = C + Ke^{rt} - K > 0$. If the buyer of the put does not exercise early, then the arbitrageur holds the portfolio until options expire, and its value becomes $C_T + Ke^{rT} - P_T - S_T = Ke^{rT} - K > 0$.

Therefore, as long as the bid and ask prices of the actual stock remain within the bounds (12) and (13), there will be no arbitrage opportunities. The actual stock price can vary within

these bounds, which means that the stock and the options markets are no longer tightly linked like in the model of An et al. (2014). In this new setting, if there is more information in options compared to stocks, then the synthetic stock price is expected to deviate from the actual stock price. Specifically, if informed traders with a positive (negative) signal are more likely to trade in options, they will cause the price of calls to increase (decrease), and the price of puts to decrease (increase), which then increases (decrease) the synthetic stock price relative to the actual price. The actual stock price will then adjust with a delay as the new information gets revealed to the stock market. This corresponds to the cases illustrated in Panels A and B of Figure 1.

IV Empirical Analysis

The theoretical discussion described above leads to a number of implications that can be tested empirically. In particular, if an option-based measure reflects incremental information, then it should predict future *actual* stock returns, but it should not predict (or predict weakly) future *synthetic* stock returns. This can be implemented by subtracting the synthetic stock return from the actual stock return, and then test if the option-based measures predict this difference. If it does, then the options market is informationally more efficient than the stock market, and there is incremental information in options. If it does not, then this would be evidence inconsistent with the greater informational efficiency of the options market, and it could also cast doubt on the incremental information reflected in option measures.

A Sample Construction

The option data, which consists of option prices, implied volatilities, open interest, and option volume, are extracted from the OptionMetrics database. The stock data, including prices, returns, volume, and the number of shares outstanding, for common stock (share

codes 10 and 11) traded on the NYSE, NASDAQ, and AMEX (exchange codes 1, 2, and 3), are extracted from CRSP. Earnings announcement dates are from I/B/E/S, and 8-K filing dates are from SEC Analytics Suite. The sample used in this paper covers the period from 1996 to 2013.

After applying a number of filters to the option data, the merging of I/B/E/S with OptionMetrics results in a final sample of about 68k earnings announcements. These are assumed to be pre-scheduled corporate events.⁷ The merging of SEC Analytics Suite with OptionMetrics results in about 85k 8-K filings that are unrelated to earnings. Only the first filing of every month is included in the sample, and only if it is not within a week before or after an earnings announcement date. These are all considered to be unscheduled corporate events.

Following Gao and Ritter (2010), reported volume data for stocks trading on NASDAQ is adjusted in the period before 2004. The stock data is merged with daily options data from OptionMetrics. Pairs of call and put options with the same maturities and strike prices are formed whenever both options satisfy the following restrictions: (i) their expiration date is at least 10 days but not more than 30 days away, (ii) their annualized implied volatility does not exceed 250%, (iii) their bid prices are non-missing and are strictly positive, (iv) their ask prices are non-missing and are strictly greater than their bid prices, (v) their open interest is greater than zero, and (vi) their deltas are non-missing. We delete option pairs for which the difference between the call delta and the put delta falls outside the interval $[0.9, 1.1]$. While not explicitly designed to do so, these option filters remove penny stocks (i.e., stocks with closing prices below \$5 (\$1) before (after) April 2001) from the final sample.

To compute the option-implied bounds in (12) and (13), continuously compounded risk-

⁷ It is typical to classify earnings announcement dates as pre-scheduled and non-discretionary since the SEC requires firms to report earnings within 35 (60) days after quarter (year) end. Prior research finds little flexibility: Bagnoli et al. (2002) find that only 1.8% (1.6%) of firms are more than 7 days late (early) compared with their pre-announced release dates, and that being late leads to a negative market reaction.

free rates are extracted from the OptionMetrics Zero Coupon Yield Curve dataset and are used to calculate the present value of the strike price and the present value of dividends with ex-dates prior to the maturity of the options in each put-call pair. The present value of each dividend is calculated by discounting back from its payment date.⁸ A firm is defined to be a dividend payer if its distribution in the OptionMetrics Dividend dataset is of type 1 (i.e., cash dividends) as this is most likely to affect the option-implied stock price bounds.⁹

The options in the sample are written on individual stocks and are generally of the American type. The synthetic (option-implied) stock price is taken to be the mid-point of the lower and the upper bounds, as defined in (12) and (13). The synthetic stock return is computed as the percent change in the synthetic stock price. To minimize the noise in the estimation, synthetic stock prices are computed by aggregating across all put-call pairs using the inverse of option bid-ask spreads in each pair as weight. This assigns a lower weight to pairs of options with wider bid-ask spreads. This follows the procedure used in Goncalves-Pinto et al. (2020).¹⁰

The stock returns used in this empirical analysis are mid-quote returns, i.e., the percent change in stock mid-quotes. To account for the effect of firm size, the reported portfolio returns are all value-weighted, using stock-level market capitalization as weight.

B Empirical Results

The period around the release of material corporate news provides investors stronger incentives to gather private information and trade on such information. These periods are then

⁸ The standardized set of maturities for risk-free rates in OptionMetrics is linearly interpolated for other maturities. The realized dividends are then discounted at the risk-free rate. The dividends in this sample are almost certain since they have almost always been announced and are payable within 22 calendar days.

⁹ The final sample excludes all stocks with a liquidating dividend (LIQUID_FLAG=1), a dividend cancellation (CANCEL_FLAG=1), a stock dividend, a stock split or a special dividend (DISTR_TYPE = 2, 3, or 5, with AMOUNT > 0) before the option expiration date, as these events may give rise to adjustments in the terms of the option contracts.

¹⁰ The results would be qualitatively similar if the aggregation across pairs of options on a given stock-date was done using the open interest of the pair as weight, like in Cremers and Weinbaum (2010).

typically used to validate informed trading proxies, like in Amin and Lee (1997), Roll et al. (2010), Johnson and So (2012), among others.

The empirical analysis in this paper focuses primarily on the period around scheduled earnings announcements. The results for unscheduled non-earnings 8-K filings are reported in Section IV.B.6.

B.1 Scheduled News and Option Trading: Anticipation or Processing?

Table I reports the predictability derived from the option to stock volume (O/S) ratio of Roll et al. (2010), around scheduled earnings announcements. This measure has been used in prior research as a proxy for informed option trading. For instance, Johnson and So (2012) show that O/S predicts future firm-specific earnings news, and argue that this is consistent with the idea that this measure reflects private information.

[Insert Table I about here]

Panel A shows that O/S on the day before an earnings announcement ($t = -1$) is a strong predictor of actual stock returns during the announcement window.¹¹ The actual returns are computed as percent changes in mid-quote prices, and reported in rows (1) to (3).

The long-short portfolio is constructed as follows. First, every quarter, ten groups of stocks are formed by sorting based on the values of O/S on the day before the earnings announcement. Second, the portfolio goes long on stocks with high O/S values (top decile), and goes short on the stocks with low O/S values (bottom decile). The cumulative actual return of the long-short portfolio during the announcement window (i.e., the announcement day plus the two subsequent days $[0, 2]$ in row (2)) is -66 basis points (bps). If the announcement day and the day after are skipped (i.e., the analysis focuses on the window $[2, 5]$

¹¹ The time line used in this section is different from the one used in Figure 1. An earnings announcement day is denoted as $t = 0$, and the day before is denoted as $t = -1$. In Figure 1, the announcement date, which is the date in which the information is revealed and becomes public, corresponds to $t = 2$.

in row (3)), then the cumulative actual return of the long-short portfolio is -52 bps. These results are statistically significant at the 1% level.

This measure consistently predicts returns negatively. Johnson and So (2012) argue that this could be because informed investors prefer to trade options when holding negative signals. However, the results show that most of the predictive ability of O/S around earnings releases is derived from stocks with low O/S values rather than the stocks with high O/S values.

Rows (4) to (6) of Panel A show that the synthetic returns of the O/S long-short portfolio formed on the day before the earnings announcement date are very similar to those using actual returns. Specifically, for the window $[0, 2]$ the synthetic long-short return is -50 bps (i.e., row (5)), and for the window $[2, 5]$ it is -49 bps (i.e., row (6)). These returns are also statistically significant at the 1% level.

The difference between actual and synthetic long-short returns is not significantly different from zero in either of the two windows (i.e., the difference between rows (2) and (5), and the difference between rows (3) and (6), respectively). Following the implications of the model described above, these results are inconsistent with the idea that the options market is informationally more efficient than the stock market, and that the relative trading activity in options versus stocks conveys any incremental information that is not already reflected in stocks. In fact, most of the predictive ability of O/S before earnings announcements is derived from the stocks with low O/S values.

Panel B examines the predictive ability of O/S on the days of earnings announcements ($t = 0$). This test captures whether option traders are good at processing publicly available information, following Engelberg et al. (2012) and Cremers et al. (2018). The procedure for conducting this test is as above, except that portfolios are formed on the day of the announcement, instead of the day before. The results show that O/S is also a strong predictor of actual and synthetic long-short returns when portfolios are formed on earnings days.

Regarding the difference between actual and synthetic long-short returns, it is not statistically different from zero for the three days after the announcement (i.e., the window $[1, 3]$), but it is negative and significant at the 5% level for the window $[2, 5]$. This suggests that option traders may be better at processing information once it is made public, but they do not appear to have an advantage in the anticipation of earnings releases (Panel A).

B.2 Option Leverage

The rational model of An et al. (2014) described in Section III predicts that in equilibrium the informed investor holding a positive signal is more likely to purchase cheap (out-of-the-money) call options and write expensive (in-the-money) put options. To account for the effect of leverage on the predictive ability of option volume, this paper repeats the analysis described above using the volume-weighted strike-stock price (VWKS) ratio of Bernile et al. (2017). It is computed as follows:

$$VWKS_{i,t} = \frac{\sum_{j=1}^n Volume_{i,t,j} \times (\frac{K_{i,t,j}}{S_{i,t}} - 1)}{\sum_{j=1}^n Volume_{i,t,j}} \quad (14)$$

where $K_{i,t,j}$ is the strike price of contract j , for stock i and day t , $Volume_{i,t,j}$ is the trading volume of contract j , n is the total number of unique option contracts available, and $S_{i,t}$ is the underlying stock price.

A high (low) value of VWKS indicates that the option volume distribution across strikes is skewed towards the out-of-the-money call (put) options. Therefore, it suggests that informed investors holding a positive (negative) signal are looking to leverage up their private information by purchasing out-of-the-money calls (puts). This measure is expected to predict returns of the underlying stock positively.

Table II reports the returns of the long-short VWKS portfolio formed on the day before earnings announcement days (Panel A). This predictor variable generates an actual (mid-quote) return of 94 bps for the window $[0, 2]$, significant at the 1% level. For the window

[2, 5], the return is 51 bps, significant only at the 10% level. The reason why the results are stronger for VWKS than O/S in [0, 2] but weaker in [2, 5], is that the VWKS measure includes the underlying stock price in the computation of the moneyness ratio (K/S), and is therefore more sensitive to stock price pressure, as argued in Goncalves-Pinto et al. (2020).¹²

[Insert Table II about here]

The difference between actual and synthetic long-short VWKS returns is 5 bps for [0, 2] and 7 bps for [2, 5], neither of which is statistically different from zero. Similar to using O/S, the results using VWKS are also inconsistent with the greater informational efficiency of the options market.

If the VWKS long-short portfolios are formed on the announcement day (Panel B) instead of the day before (Panel A), the difference between actual and synthetic long-short returns is -15 bps significant at the 10% level for the window [0, 2], and -3 bps not significantly different from zero for the window [2, 5]. The negative difference for [0, 2] is also inconsistent with option traders being superior information processors, because the predictability of synthetic prices is larger than that of actual prices. This result does not follow the rationale of Figure 1. Instead, in Panel B the actual stock return on the announcement day ($t = 0$) is larger than the synthetic return. This means that actual prices lead synthetic prices, and suggests that stock traders are better processors of public information.

B.3 How Interconnected Are Stock and Options Markets?

Panel C of Figure 1 suggests that the failure of option measures to predict the difference between actual and synthetic returns could be driven by the tight interconnectedness between stock and options markets. If the two markets are tightly connected, this does not allow synthetic prices to deviate significantly from actual stock prices. This would be the case if

¹² Section IV.B.5 discusses the issue of stock price pressure in more detail.

the option contracts are European-type, in which case the prices of stock and options are tightly linked by put-call parity, and any deviation from that relation leads to an arbitrage opportunity.

However, the options on individual stocks that are used in this empirical analysis are typically of the American-type, and option bid-ask spreads are quite large. This means that the put-call parity relation is a pair of inequalities (Cox and Rubinstein (1985)). This results in a no-arbitrage price range within which the stock price can vary without leading to arbitrage opportunities.

Table III reports the difference between the upper and the lower no-arbitrage bounds, as a percent of the mid-point, for the portfolios of stocks created in Tables I and II. The no-arbitrage price range, as a percent of the mid-point, is defined as follows: $NAPR = (S^U - S^L)/MP$, where the mid-point is $MP = (S^U + S^L)/2$, and the bounds S^L and S^U are defined in (12) and (13).

[Insert Table III about here]

Panel A reports the results for the portfolios formed on the the day before the earnings announcement ($t = -1$). It shows the $NAPR$ not only for the formation date but also the days surrounding it. Specifically, it shows the results for days $t \in \{-3, -2, -1, 0, 1, 2, 3\}$, where day 0 is the day of the earnings announcement, and day -1 is the portfolio formation date. The results show that $NAPR$ is larger on the day of portfolio formation, but it continues to be large on the surrounding days. If O/S is used to form portfolios, the values of $NAPR$ range from 1.37% to 1.46% for the portfolios in the high O/S group, and ranges from 2.11% to 2.40% for the low O/S group. If VWKS is used to form portfolios, the values of $NAPR$ range from 1.44% to 1.60% for the portfolios in the high VWKS group, and ranges from 2.02% to 2.20% for the low VWKS group.

The patterns are similar when forming portfolios on the day of the earnings announcement (Panel B).

This suggests that synthetic and actual stock prices deviate from each other significantly without leading to arbitrage opportunities. Therefore, the linkage between the two markets is not tight, and this cannot explain the results reported in Tables I and II.

B.4 Comparison with Option Price-Based Predictors

Table IV reports the predictability results around earnings news using two sorting variables based on option prices: (i) the distance between the synthetic and the actual stock price as a percent of the actual price (DOTS), and (ii) the implied-volatility spread (IVS). The first measure is from Goncalves-Pinto et al. (2020), and the second is from Cremers and Weinbaum (2010). It can be argued that, under some assumptions, DOTS and IVS are essentially the same measure, but the first measure has the advantage of being model-free. These two measures, and any other measure that depends on the traded stock price (like the VWKS measure described above) are very sensitive to price pressure in the stock market. The results reported in Table IV are indeed consistent with such strong sensitivity to stock price pressure.

[Insert Table IV about here]

The results show that, both DOTS and IVS predict actual stock returns, but they do not predict synthetic stock returns. This appears to be consistent with Panel A of Figure 1. However, a closer look at the results reveals that this predictability is driven by the strong return reversals observed from the formation day ($t = -1$) to the holding period. This means that stock price pressure is the likely primary driver of these results.¹³

The two measures strongly predict the difference between actual and synthetic long-short returns for the window $[0, 2]$, but after accounting for the effect of stock price pressure (i.e.,

¹³ Goncalves-Pinto et al. (2020) also examine the predictability of DOTS around earnings announcements. However, their results are quantitatively different because Table IV (Panel A) reports mid-quote returns, while they report traded stock returns.

skipping a couple of days between the formation and the holding periods), the predictability vanishes for the window $[2, 5]$. It is also clear that there is negative autocorrelation in returns when comparing the results for the formation day ($t = -1$) and the holding period $[0, 2]$, consistent with stock price pressure.

B.5 Contribution of Stock Price Pressure to the Predictability of Option Volume-Based Measures

Table V examines the impact of stock price pressure on the predictive ability of the option volume-based measures. It uses the DOTS measure from Goncalves-Pinto et al. (2020) as a proxy for stock price pressure. Then, it performs a sequential double sort based on the DOTS measure first, and within each DOTS group, sort stocks based on either O/S or VWKS. It uses terciles in both sorts, which results in 9 portfolios. The table only reports the top and bottom terciles of this sequential double sort.

[Insert Table V about here]

Panel A uses O/S as the second sorting variable. The results of this conditional analysis show that the predictive ability of O/S does not appear to be sensitive to stock price pressure. It is a strong predictor of returns within both the high and low DOTS groups.

Panel B uses VKWS as the second sorting variable. In contrast to the results in Panel A, the ability of VWKS to predict future stock returns only works for the high DOTS group, and only for the window $[0, 2]$. It fails to predict both actual and synthetic stock returns when skipping a couple of days between formation and holding (i.e., window $[2, 5]$). This is consistent with VWKS being highly sensitive to stock price pressure, which is expected given that this measure is a function of actual stock prices.

More importantly, the data fails to reject the hypothesis that actual and synthetic long-short returns are equal, in either the low or the high DOTS groups, for both predictors O/S

and VWKS. This is consistent with the conclusion from Panel A of Tables I and II, that the options market does not appear to be informationally more efficient than the stock market.

B.6 Unscheduled Corporate Events

The discussion in previous sections focuses on the period around earnings announcements. This section extends the analysis to the period around the disclosure of (non-earnings) 8-K filings, extracted from SEC Analytics Suite. This later group of events is generally not pre-scheduled, unlike earnings announcements. Therefore, the dynamics of price discovery across stock and options markets can be different around such unscheduled events.

There exist alternative sources of unscheduled corporate events. For instance, Jin et al. (2012) use unscheduled events from the Capital IQ Key Developments database, Cremers et al. (2018) use news from Thomson Reuters News Analytics, and Augustin et al. (2016) use Dow Jones news from RavenPack. The main advantage of using the 8-K filings is that it covers a much longer time period, including the entire period covered by OptionMetrics. The alternative databases only cover the more recent period.

Table VI replicates the analysis in Table I for the period around the release of (non-earnings) 8-K filings. The final sample of 8-K filings includes only the first filing of every month, and only if it is not within a week before or after an earnings announcement date.

[Insert Table VI about here]

Overall, the results in Table VI are much weaker compared to the results in Table I, but they are qualitatively very similar: O/S is a good predictor of future stock returns, both when forming portfolios on the day before the disclose of the 8-K filings (Panel A), and when forming portfolios on the disclosure date (Panel B). The predictability is stronger in Panel B, which is consistent with the idea that it is much harder to anticipate the release of unscheduled 8-K filings, compared to waiting for such news to be released first so that one can then process them.

More importantly, the null hypothesis that actual and synthetic long-short returns are equal cannot be rejected in the empirical tests of Table VI, like in the case of earnings announcements in Table I. This stresses the idea that the options market does not appear to be informationally more efficient than the stock market, and casts doubt in the existence of incremental information in options.

V Conclusion

Much of prior research argues that informed investors prefer to trade in the options market and this explains why option-based measures are strong predictors of future stock returns.

If informed investors do indeed prefer to trade in the options market, and stock and options markets are not tightly linked because of market frictions, then the synthetic (option-implied) stock price should adjust towards the fundamental stock value to a larger extent than the actual stock price. Therefore, one would be able to use option measures to predict future *actual* stock returns, but there should be weaker or no predictability for future *synthetic* stock returns, depending on the extent of noise trading in options. To the best of our knowledge, this latter condition has never been examined in prior research and we do this along three dimensions. First, we introduce a new method to determine if options contain information not yet reflected into stock prices, by focusing on the predictability of the difference between actual and synthetic stock returns. Second, we find that existing proxies for informed option trading, such as the option-to-stock volume ratio, predict both actual and synthetic stock returns to the same extent, around the release of scheduled and unscheduled firm-specific news. This is inconsistent with the idea that options convey incremental information that is not already reflected in stocks. Lastly, we motivate the empirical approach using a noisy rational expectations model with informed investors who can trade simultaneously in stock and options, following An et al. (2014).

References

- Amin, K. and C. Lee (1997). Option trading, price discovery, and earnings news dissemination. *Contemporary Accounting Research* 14, 153–192.
- An, B.-J., A. Ang, T. G. Bali, and N. Cakici (2014). The joint cross-section of stocks and options. *Journal of Finance* 69(5), 2279–2337.
- Augustin, P., M. Brenner, G. Grass, and M. G. Subrahmanyam (2016). How do insiders trade? Technical report.
- Back, K. (1993). Asymmetric information and options. *Review of Financial Studies* 6, 435–472.
- Bagnoli, M., W. Kross, and S. Watts (2002). The information in management’s expected earnings report date: A day late, a penny short. *Journal of Accounting Research* 40, 1275–96.
- Bernile, G., F. Gao, and J. Hu (2017). Center of volume mass: Does aggregate options market opinion predict future equity returns? Technical report.
- Biais, B. and P. Hillion (1994). Insider and liquidity trading in stock and options markets. *Review of Financial Studies* 7, 743–780.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 637–654.
- Cao, J., B. Han, Q. Tong, and X. Zhan (2018). Option return predictability. Technical report.
- Chabakauri, G., K. Yuan, and K. Zachariadis (2017). Multi-asset noisy rational expectations equilibrium with contingent claims. Technical report.
- Chan, K., Y. Chung, and W. Fong (2002). The informational role of stock and option volume. *Review of Financial Studies* 15, 1049–1075.
- Chan, K., Y. Chung, and H. Johnson (1993). Why option prices lag stock prices: A trading based explanation. *Journal of Finance* 48, 1957–1967.

- Cherian, J. (1993). Option pricing, self-fulfilling prophecies, implied volatilities, and strategic interaction. Technical report.
- Collin-Dufresne, P., V. Fos, and D. Muravyev (2017). Informed trading in the stock market and option price discovery. Technical report.
- Cox, J. C. and M. Rubinstein (1985). Options markets. *Prentice Hall*.
- Cremers, M., A. Fodor, D. Muravyev, and D. Weinbaum (2018). How do informed option traders trade? option trading activity, news releases, and stock return predictability. Technical report.
- Cremers, M. and D. Weinbaum (2010). Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis* 45(2), 335–367.
- Easley, D., M. O’Hara, and P. S. Srinivas (1998). Option volume and stock prices: Evidence on where informed traders trade. *Journal of Finance* 53(2), 431–465.
- Engelberg, J., A. Reed, and M. Ringgenberg (2012). How are shorts informed? short sellers, news, and information processing. *Journal of Financial Economics* 105, 260–278.
- Goncalves-Pinto, L., B. D. Grundy, A. Hameed, T. van der Heijden, and Y. Zhu (2020). Why do option prices predict stock returns? the role of price pressure in the stock market. *Management Science* 16, 3903–3926.
- Grossman, S. and J. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review* 70, 393–408.
- Hu, J. (2014). Does option trading convey stock price information? *Journal of Financial Economics* 111(3), 625–645.
- Jin, W., J. Livnat, and Y. Zhang (2012). Option prices leading equity prices: Do option traders have an information advantage? *Journal of Accounting Research* 50, 401–432.
- Johnson, T. and E. So (2012). The option to stock volume ratio and future returns. *Journal of Financial Economics* 106(2), 262–286.

- Lamont, O. A. and R. H. Thaler (2003). Can the market add and subtract? mispricing in tech stock carve-outs. *Journal of Political Economy* 111, 227–268.
- Muravyev, D., N. Pearson, and J. Broussard (2013). Is there price discovery in equity options? *Journal of Financial Economics* 107, 259–283.
- Newey, W. K. and K. D. West (1987). A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Ofek, E., M. Richardson, and R. F. Whitelaw (2004). Limited arbitrage and short sales restrictions: Evidence from the options markets. *Journal of Financial Economics* 74(2), 305–342.
- Pan, J. and A. M. Poteshman (2006). The information in option volume for future stock prices. *Review of Financial Studies* 19(3), 871–908.
- Roll, R., E. Schwartz, and A. Subrahmanyam (2010). O/s: The relative trading activity in options and stock. *Journal of Financial Economics* 96, 1–17.
- Stephan, J. A. and R. E. Whaley (1990). Intraday price change and trading volume relations in the stock and stock option markets. *Journal of Finance* 45(1), 191–220.
- Vijh, A. (1990). Liquidity of the cboe equity options. *Journal of Finance* 45, 1157–1179.
- Xing, Y., X. Zhang, and R. Zhao (2010). What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis* 45, 641–662.

Table I: Predictability of O/S Around Earnings Announcements

This table reports the returns of portfolios of stocks formed based on O/S, the option to stock volume ratio of Roll et al. (2010), around quarterly earnings announcements. In Panel A, stocks are sorted into deciles based on O/S on the day before the announcement ($t = -1$), and in Panel B the portfolios are formed using the O/S on the announcement day ($t = 0$). Rows (1) to (3) report value-weighted mid-quote stock returns in excess of the risk-free rate, and rows (4) to (6) report value-weighted synthetic stock returns in excess of the risk-free rate. Synthetic returns are computed as percent changes in synthetic stock prices, which in turn are computed as the mid-point of the upper and lower no-arbitrage bounds defined in Section III.D.2. The table shows the results for the portfolios of stocks in the top and bottom O/S deciles (“High O/S” and “Low O/S” respectively), as well as the results of the long-short portfolio that purchases the “High O/S” group and shorts the “Low O/S” group (i.e., “High-Low”). The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Single sort on O/S before earnings days ($t = -1$) | | | | | |
|---|-----------|--------|---------------------|---------------------|----------------------|
| Return | Row | Period | Low O/S | High O/S | High - Low |
| Actual | (1) | -1 | 0.199*** (0.049) | 0.318*** (0.115) | 0.119 (0.125) |
| Actual | (2) | [0,2] | 0.558*** (0.134) | -0.103 (0.159) | -0.661*** (0.190) |
| Actual | (3) | [2,5] | 0.342*** (0.117) | -0.177 (0.138) | -0.518*** (0.166) |
| Synthetic | (4) | -1 | 0.127*** (0.040) | 0.303*** (0.110) | 0.175 (0.109) |
| Synthetic | (5) | [0,2] | 0.336*** (0.110) | -0.167 (0.141) | -0.503*** (0.165) |
| Synthetic | (6) | [2,5] | 0.260** (0.103) | -0.232* (0.138) | -0.492*** (0.158) |
| Actual - Synthetic | (1) - (4) | -1 | 0.071*** (0.026) | 0.015 (0.029) | -0.057 (0.035) |
| Actual - Synthetic | (2) - (5) | [0,2] | 0.223*** (0.069) | 0.065 (0.082) | -0.158 (0.101) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.081* (0.044) | 0.055* (0.031) | -0.026 (0.052) |
| Panel B: Single sort on O/S on earnings days ($t = 0$) | | | | | |
| Return | Row | Period | Low O/S | High O/S | High - Low |
| Actual | (1) | 0 | 0.168* (0.100) | 0.426*** (0.078) | 0.258** (0.122) |
| Actual | (2) | [1,3] | 0.523*** (0.132) | -0.365 (0.238) | -0.888*** (0.289) |
| Actual | (3) | [2,5] | 0.366*** (0.105) | -0.359** (0.158) | -0.724*** (0.167) |
| Synthetic | (4) | 0 | 0.064 (0.077) | 0.392*** (0.074) | 0.327*** (0.123) |
| Synthetic | (5) | [1,3] | 0.407*** (0.101) | -0.460** (0.186) | -0.867*** (0.217) |
| Synthetic | (6) | [2,5] | 0.281*** (0.088) | -0.264* (0.140) | -0.545*** (0.139) |
| Actual - Synthetic | (1) - (4) | 0 | 0.104** (0.050) | 0.035 (0.039) | -0.069 (0.059) |
| Actual - Synthetic | (2) - (5) | [1,3] | 0.116** (0.048) | 0.095 (0.086) | -0.022 (0.103) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.084* (0.048) | -0.095 (0.065) | -0.179** (0.085) |

Table II: Predictability of VWKS Around Earnings Announcements

This table reports the returns of portfolios of stocks formed based on VWKS, the volume-weighted strike-spot ratio of Bernile et al. (2017), around quarterly earnings announcements. In Panel A, stocks are sorted into deciles based on VWKS on the day before the announcement ($t = -1$), and in Panel B the portfolios are formed using the VWKS on the announcement day ($t = 0$). Rows (1) to (3) report value-weighted mid-quote stock returns in excess of the risk-free rate, and rows (4) to (6) report value-weighted synthetic stock returns in excess of the risk-free rate. Synthetic returns are computed as percent changes in synthetic stock prices, which in turn are computed as the mid-point of the upper and lower no-arbitrage bounds defined in Section III.D.2. The table shows the results for the portfolios of stocks in the top and bottom VWKS deciles (“High VWKS” and “Low VWKS” respectively), as well as the results of the long-short portfolio that purchases the “High VWKS” group and shorts the “Low VWKS” group (i.e., “High-Low”). The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Single sort on VWKS before earnings days ($t = -1$) | | | | | |
|--|-----------|--------|---------------------|---------------------|----------------------|
| Return | Row | Period | Low VWKS | High VWKS | High - Low |
| Actual | (1) | -1 | 0.858*** (0.131) | 0.108 (0.094) | -0.750*** (0.146) |
| Actual | (2) | [0,2] | -0.011 (0.169) | 0.930*** (0.156) | 0.941*** (0.233) |
| Actual | (3) | [2,5] | -0.138 (0.136) | 0.374* (0.198) | 0.512* (0.277) |
| Synthetic | (4) | -1 | 0.785*** (0.115) | 0.118 (0.090) | -0.666*** (0.106) |
| Synthetic | (5) | [0,2] | -0.134 (0.179) | 0.757*** (0.141) | 0.890*** (0.210) |
| Synthetic | (6) | [2,5] | -0.109 (0.137) | 0.334* (0.177) | 0.442* (0.264) |
| Actual - Synthetic | (1) - (4) | -1 | 0.073 (0.056) | -0.010 (0.051) | -0.083 (0.082) |
| Actual - Synthetic | (2) - (5) | [0,2] | 0.122* (0.067) | 0.173** (0.084) | 0.051 (0.107) |
| Actual - Synthetic | (3) - (6) | [2,5] | -0.029 (0.060) | 0.041 (0.093) | 0.070 (0.111) |

| Panel B: Single sort on VWKS on earnings days ($t = 0$) | | | | | |
|---|-----------|--------|---------------------|---------------------|----------------------|
| Return | Row | Period | Low VWKS | High VWKS | High - Low |
| Actual | (1) | 0 | 2.213*** (0.197) | 0.358*** (0.096) | -1.854*** (0.197) |
| Actual | (2) | [1,3] | -0.158 (0.206) | 0.673*** (0.178) | 0.830*** (0.274) |
| Actual | (3) | [2,5] | -0.210 (0.177) | 0.316** (0.154) | 0.526*** (0.195) |
| Synthetic | (4) | 0 | 2.096*** (0.209) | 0.238*** (0.085) | -1.858*** (0.198) |
| Synthetic | (5) | [1,3] | -0.371* (0.201) | 0.610*** (0.170) | 0.981*** (0.243) |
| Synthetic | (6) | [2,5] | -0.246* (0.130) | 0.307** (0.132) | 0.553*** (0.187) |
| Actual - Synthetic | (1) - (4) | 0 | 0.117 (0.082) | 0.121* (0.062) | 0.004 (0.078) |
| Actual - Synthetic | (2) - (5) | [1,3] | 0.213*** (0.073) | 0.062 (0.051) | -0.151* (0.087) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.036 (0.073) | 0.009 (0.100) | -0.027 (0.090) |

Table III: No-Arbitrage Price Range Around Earnings Announcements

This table reports the difference between the upper and the lower no-arbitrage bounds, as a percent of the mid-point, for the portfolios of stocks created in Tables I and II. The no-arbitrage price range, as a percent of the mid-point, is defined as follows: $NAPR = (S^U - S^L)/MP$, where the mid-point is $MP = (S^U + S^L)/2$, and the bounds S^L and S^U are defined in Section III.D.2. Panel A reports the results for the portfolios formed on the the day before the earnings announcement ($t = -1$), and Panel B reports the results for the portfolios formed on the announcement day ($t = 0$). It shows the $NAPR$ not only for the formation date but also the days surrounding it. Specifically, it shows the results for days $t \in \{-3, -2, -1, 0, 1, 2, 3\}$, where day 0 is the day of the earnings announcement. The table shows the results for the portfolios of stocks in the top and bottom deciles for each of the sorting variables, as well as the results of the long-short portfolio (“High-Low”). The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Single sort before earnings days (t = -1) | | | | | | |
|--|---------------------|---------------------|----------------------|---------------------|---------------------|----------------------|
| Period | Sort variable: O/S | | | Sort variable: VWKS | | |
| | Low O/S | High O/S | High - Low | Low VWKS | High VWKS | High - Low |
| -3 | 2.246*** (0.119) | 1.462*** (0.109) | -0.784*** (0.160) | 2.177*** (0.088) | 1.561*** (0.078) | -0.616*** (0.056) |
| -2 | 2.283*** (0.117) | 1.460*** (0.112) | -0.824*** (0.160) | 2.164*** (0.080) | 1.569*** (0.077) | -0.595*** (0.057) |
| -1 | 2.395*** (0.133) | 1.462*** (0.108) | -0.932*** (0.170) | 2.202*** (0.075) | 1.595*** (0.076) | -0.607*** (0.049) |
| 0 | 2.333*** (0.127) | 1.458*** (0.109) | -0.874*** (0.160) | 2.169*** (0.077) | 1.580*** (0.077) | -0.589*** (0.044) |
| 1 | 2.218*** (0.105) | 1.370*** (0.116) | -0.847*** (0.140) | 2.049*** (0.078) | 1.473*** (0.079) | -0.576*** (0.041) |
| 2 | 2.165*** (0.107) | 1.375*** (0.117) | -0.790*** (0.142) | 2.015*** (0.089) | 1.457*** (0.080) | -0.558*** (0.041) |
| 3 | 2.113*** (0.106) | 1.366*** (0.112) | -0.747*** (0.146) | 2.016*** (0.080) | 1.436*** (0.083) | -0.579*** (0.039) |

| Panel B: Single sort on earnings days (t = 0) | | | | | | |
|---|---------------------|---------------------|----------------------|---------------------|---------------------|----------------------|
| Period | Sort variable: O/S | | | Sort variable: VWKS | | |
| | Low O/S | High O/S | High - Low | Low VWKS | High VWKS | High - Low |
| -3 | 2.322*** (0.152) | 1.488*** (0.106) | -0.834*** (0.184) | 2.225*** (0.077) | 1.554*** (0.072) | -0.671*** (0.052) |
| -2 | 2.346*** (0.144) | 1.492*** (0.106) | -0.854*** (0.181) | 2.267*** (0.072) | 1.577*** (0.076) | -0.690*** (0.055) |
| -1 | 2.427*** (0.152) | 1.484*** (0.107) | -0.943*** (0.183) | 2.284*** (0.072) | 1.595*** (0.077) | -0.689*** (0.056) |
| 0 | 2.502*** (0.159) | 1.528*** (0.105) | -0.974*** (0.194) | 2.323*** (0.069) | 1.625*** (0.074) | -0.698*** (0.059) |
| 1 | 2.322*** (0.135) | 1.401*** (0.112) | -0.922*** (0.171) | 2.136*** (0.073) | 1.503*** (0.078) | -0.633*** (0.037) |
| 2 | 2.305*** (0.130) | 1.372*** (0.113) | -0.933*** (0.166) | 2.073*** (0.073) | 1.469*** (0.080) | -0.603*** (0.036) |
| 3 | 2.282*** (0.136) | 1.360*** (0.111) | -0.922*** (0.176) | 2.085*** (0.072) | 1.449*** (0.080) | -0.636*** (0.036) |

Table IV: Predictability of DOTS and IVS

This table reports the predictability results around earnings announcements using two sorting variables based on option prices. Panel A uses the distance between the synthetic and the actual stock price as a percent of the actual price (DOTS), and Panel B uses the implied-volatility spread (IVS). The first measure is from Goncalves-Pinto et al. (2020), and the second is from Cremers and Weinbaum (2010). In both panels, portfolios are formed on the day before the announcement ($t = -1$). Rows (1) to (3) report value-weighted mid-quote stock returns in excess of the risk-free rate, and rows (4) to (6) report value-weighted synthetic stock returns in excess of the risk-free rate. Synthetic returns are computed as percent changes in synthetic stock prices, which in turn are computed as the mid-point of the upper and lower no-arbitrage bounds defined in Section III.D.2. The table shows the results for the portfolios of stocks in the top and bottom deciles of the sorting variables, as well as the results of the long-short portfolio (i.e., “High-Low”). The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Single sort on DOTS before earnings days ($t = -1$) | | | | | |
|--|-----------|--------|----------------------|----------------------|----------------------|
| Return | Row | Period | Low DOTS | High DOTS | High - Low |
| Actual | (1) | -1 | 0.665*** (0.114) | -0.367** (0.158) | -1.032*** (0.226) |
| Actual | (2) | [0,2] | 0.000 (0.189) | 0.514** (0.240) | 0.515* (0.259) |
| Actual | (3) | [2,5] | 0.194 (0.225) | 0.060 (0.196) | -0.135 (0.222) |
| Synthetic | (4) | -1 | 0.376*** (0.098) | -0.166 (0.150) | -0.542** (0.208) |
| Synthetic | (5) | [0,2] | 0.344* (0.201) | 0.271 (0.253) | -0.073 (0.263) |
| Synthetic | (6) | [2,5] | 0.177 (0.176) | 0.093 (0.182) | -0.084 (0.193) |
| Actual - Synthetic | (1) - (4) | -1 | 0.289*** (0.037) | -0.201*** (0.041) | -0.490*** (0.062) |
| Actual - Synthetic | (2) - (5) | [0,2] | -0.344*** (0.074) | 0.243*** (0.076) | 0.587*** (0.103) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.017 (0.082) | -0.033 (0.078) | -0.051 (0.103) |
| Panel B: Single sort on IVS before earnings days ($t = -1$) | | | | | |
| Return | Row | Period | Low IVS | High IVS | High - Low |
| Actual | (1) | -1 | 0.645*** (0.150) | -0.335** (0.164) | -0.980*** (0.260) |
| Actual | (2) | [0,2] | -0.050 (0.203) | 0.382* (0.201) | 0.433* (0.232) |
| Actual | (3) | [2,5] | 0.081 (0.208) | 0.208 (0.172) | 0.128 (0.206) |
| Synthetic | (4) | -1 | 0.371*** (0.116) | -0.090 (0.134) | -0.460** (0.189) |
| Synthetic | (5) | [0,2] | 0.093 (0.173) | 0.185 (0.203) | 0.092 (0.222) |
| Synthetic | (6) | [2,5] | 0.147 (0.181) | 0.164 (0.152) | 0.017 (0.193) |
| Actual - Synthetic | (1) - (4) | -1 | 0.275*** (0.062) | -0.245*** (0.067) | -0.520*** (0.106) |
| Actual - Synthetic | (2) - (5) | [0,2] | -0.143* (0.075) | 0.197** (0.081) | 0.340*** (0.104) |
| Actual - Synthetic | (3) - (6) | [2,5] | -0.066 (0.098) | 0.044 (0.066) | 0.111 (0.115) |

Table V: Predictability of O/S and VWKS Controlling for Price Pressure

This table reports the returns of portfolios formed using a sequential double sort, according to which stock are first sorted into terciles based on DOTS, and then within each tercile of DOTS sort stocks again into terciles based on O/S (Panel A) or VWKS (Panel B), creating a total of 9 portfolios for each panel. The difference between the synthetic and the actual stock price (DOTS) is from Goncalves-Pinto et al. (2020). The option to stock volume ratio (O/S) is from Roll et al. (2010), and the volume-weighted strike-spot ratio (VWKS) is from Bernile et al. (2017). In both panels, portfolios are formed on the day before the announcement ($t = -1$). The table reports actual returns (i.e., value-weighted mid-quote stock returns in excess of the risk-free rate), synthetic returns (i.e., value-weighted synthetic stock returns in excess of the risk-free rate), and their difference. Synthetic returns are computed as percent changes in synthetic stock prices, which in turn are computed as the mid-point of the upper and lower no-arbitrage bounds defined in Section III.D.2. The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Sequential double sort on DOTS (terciles) and O/S (terciles) before earnings days (t = -1) | | | | | | | |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|----------------------|
| Period | | Actual | | Synthetic | | Actual - Synthetic | |
| | | Low DOTS | High DOTS | Low DOTS | High DOTS | Low DOTS | High DOTS |
| -1 | Low O/S | 0.470*** (0.109) | -0.194* (0.105) | 0.354*** (0.103) | -0.068 (0.100) | 0.117*** (0.027) | -0.126*** (0.018) |
| | High O/S | 0.579*** (0.117) | -0.128 (0.085) | 0.456*** (0.115) | -0.095 (0.106) | 0.124*** (0.044) | -0.033 (0.058) |
| | High O/S - Low O/S | 0.109 (0.149) | 0.067 (0.140) | 0.102 (0.137) | -0.027 (0.153) | 0.007 (0.051) | 0.094* (0.049) |
| [0,2] | Low O/S | 0.208 (0.128) | 0.523*** (0.140) | 0.277** (0.106) | 0.397*** (0.139) | -0.069 (0.051) | 0.126*** (0.045) |
| | High O/S | -0.213 (0.201) | -0.200 (0.174) | -0.102 (0.221) | -0.227 (0.185) | -0.111* (0.060) | 0.026 (0.045) |
| | High O/S - Low O/S | -0.421** (0.185) | -0.724*** (0.191) | -0.379* (0.192) | -0.624*** (0.211) | -0.042 (0.076) | -0.100* (0.056) |
| [2,5] | Low O/S | 0.384*** (0.140) | 0.313* (0.157) | 0.322** (0.140) | 0.300** (0.139) | 0.063 (0.049) | 0.013 (0.044) |
| | High O/S | 0.019 (0.165) | -0.082 (0.195) | 0.007 (0.132) | -0.190 (0.186) | 0.012 (0.095) | 0.108** (0.045) |
| | High O/S - Low O/S | -0.365*** (0.117) | -0.395** (0.169) | -0.315*** (0.107) | -0.490*** (0.143) | -0.050 (0.115) | 0.095 (0.069) |
| Panel B: Sequential double sort on DOTS (terciles) and VWKS (terciles) before earnings days (t = -1) | | | | | | | |
| Period | | Actual | | Synthetic | | Actual - Synthetic | |
| | | Low DOTS | High DOTS | Low DOTS | High DOTS | Low DOTS | High DOTS |
| -1 | Low VWKS | 1.018*** (0.165) | 0.369*** (0.107) | 0.901*** (0.159) | 0.436*** (0.105) | 0.117*** (0.031) | -0.067 (0.054) |
| | High VWKS | 0.236** (0.109) | -0.252** (0.113) | 0.108 (0.102) | -0.150 (0.099) | 0.128*** (0.022) | -0.102* (0.061) |
| | High VWKS - Low VWKS | -0.782*** (0.196) | -0.621*** (0.140) | -0.794*** (0.186) | -0.585*** (0.113) | 0.011 (0.029) | -0.036 (0.073) |
| [0,2] | Low VWKS | -0.037 (0.202) | -0.186 (0.196) | -0.007 (0.207) | -0.269 (0.189) | -0.030 (0.071) | 0.083* (0.043) |
| | High VWKS | 0.262 (0.181) | 0.842*** (0.197) | 0.375** (0.180) | 0.686*** (0.193) | -0.112** (0.049) | 0.156*** (0.056) |
| | High VWKS - Low VWKS | 0.300 (0.262) | 1.028*** (0.276) | 0.382 (0.262) | 0.955*** (0.249) | -0.082 (0.084) | 0.073 (0.078) |
| [2,5] | Low VWKS | -0.017 (0.159) | 0.041 (0.160) | -0.004 (0.144) | -0.027 (0.148) | -0.012 (0.065) | 0.068* (0.037) |
| | High VWKS | 0.253 (0.187) | 0.171 (0.216) | 0.217 (0.176) | 0.118 (0.201) | 0.036 (0.041) | 0.053 (0.043) |
| | High VWKS - Low VWKS | 0.270 (0.245) | 0.130 (0.201) | 0.222 (0.221) | 0.145 (0.194) | 0.048 (0.069) | -0.015 (0.054) |

Table VI: Predictability of O/S Around Non-Earnings 8-K Filings

This table reports the returns of portfolios of stocks formed based on O/S, the option to stock volume ratio of Roll et al. (2010), around the disclosure of (non-earnings) 8-K filings, which are extracted from SEC Analytics Suite. In Panel A, stocks are sorted into deciles based on O/S on the day before the announcement ($t = -1$), and in Panel B the portfolios are formed using the O/S on the announcement day ($t = 0$). Rows (1) to (3) report value-weighted mid-quote stock returns in excess of the risk-free rate, and rows (4) to (6) report value-weighted synthetic stock returns in excess of the risk-free rate. Synthetic returns are computed as percent changes in synthetic stock prices, which in turn are computed as the mid-point of the upper and lower no-arbitrage bounds defined in Section III.D.2. The table shows the results for the portfolios of stocks in the top and bottom O/S deciles (“High O/S” and “Low O/S” respectively), as well as the results of the long-short portfolio that purchases the “High O/S” group and shorts the “Low O/S” group (i.e., “High-Low”). The sample period is from 1996 to 2013. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively. Newey and West (1987) standard errors are reported in parenthesis.

| Panel A: Single sort on O/S before news days ($t = -1$) | | | | | |
|---|-----------|--------|-------------------|-------------------|---------------------|
| Return | Row | Period | Low O/S | High O/S | High - Low |
| Actual | (1) | -1 | -0.016 (0.045) | 0.127 (0.077) | 0.142* (0.075) |
| Actual | (2) | [0,2] | 0.090 (0.126) | -0.245 (0.183) | -0.335** (0.151) |
| Actual | (3) | [2,5] | 0.083 (0.109) | -0.192 (0.167) | -0.275* (0.142) |
| Synthetic | (4) | -1 | -0.009 (0.040) | 0.072 (0.079) | 0.081 (0.077) |
| Synthetic | (5) | [0,2] | 0.092 (0.107) | -0.166 (0.168) | -0.258 (0.170) |
| Synthetic | (6) | [2,5] | 0.041 (0.102) | -0.184 (0.153) | -0.225* (0.134) |
| Actual - Synthetic | (1) - (4) | -1 | -0.007 (0.017) | 0.055 (0.038) | 0.062 (0.038) |
| Actual - Synthetic | (2) - (5) | [0,2] | -0.002 (0.052) | -0.079 (0.071) | -0.077 (0.067) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.042 (0.054) | -0.008 (0.055) | -0.049 (0.066) |

| Panel B: Single sort on O/S on news days ($t = 0$) | | | | | |
|--|-----------|--------|---------------------|--------------------|----------------------|
| Return | Row | Period | Low O/S | High O/S | High - Low |
| Actual | (1) | 0 | -0.004 (0.044) | 0.227** (0.100) | 0.231** (0.114) |
| Actual | (2) | [1,3] | 0.284** (0.112) | -0.092 (0.141) | -0.376*** (0.142) |
| Actual | (3) | [2,5] | 0.210** (0.089) | -0.225 (0.153) | -0.435*** (0.157) |
| Synthetic | (4) | 0 | -0.018 (0.036) | 0.176* (0.102) | 0.194* (0.109) |
| Synthetic | (5) | [1,3] | 0.250*** (0.096) | -0.076 (0.139) | -0.326** (0.135) |
| Synthetic | (6) | [2,5] | 0.151* (0.086) | -0.218 (0.145) | -0.369** (0.157) |
| Actual - Synthetic | (1) - (4) | 0 | 0.014 (0.025) | 0.051 (0.054) | 0.037 (0.067) |
| Actual - Synthetic | (2) - (5) | [1,3] | 0.034 (0.039) | -0.017 (0.053) | -0.050 (0.058) |
| Actual - Synthetic | (3) - (6) | [2,5] | 0.059 (0.054) | -0.007 (0.038) | -0.066 (0.059) |

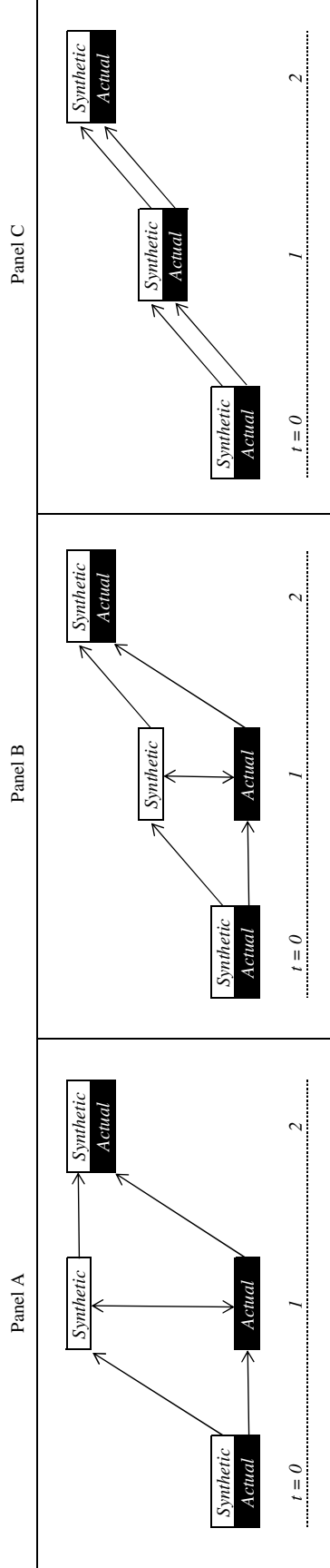


Figure 1: Actual and Synthetic Stock Prices and a Positive Private Signal

This figure illustrates the interaction between *actual* and *synthetic* stock prices when there is a positive private signal to be traded on. Panel A shows the case in which the informed investor prefers to trade in options, and noise trading is absent from the options market. To take advantage of a positive signal received at time $t = 1$, the informed investor purchases call options and sells put options. This pushes up the synthetic (option-implied) stock price immediately to the fundamental value of the stock. As the new information gets revealed to the stock traders, the actual stock price adjusts to the fundamental value at time $t = 2$. Panel B illustrates the case with the presence of noise traders in the options market. In this case, prices only adjust partially towards their efficient levels. Panels A and B assume that the stock and options markets are not fully interconnected. Panel C illustrates the case in which the two markets are fully interconnected, informed investors can trade in both markets, and there exist noise traders in both markets. In this case, synthetic and actual stock prices are always equal to each other, and there is no information in options that is not simultaneously reflected in the stock. By construction, in Panel C the predictability of actual stock returns is exactly the same as that of synthetic stock returns.

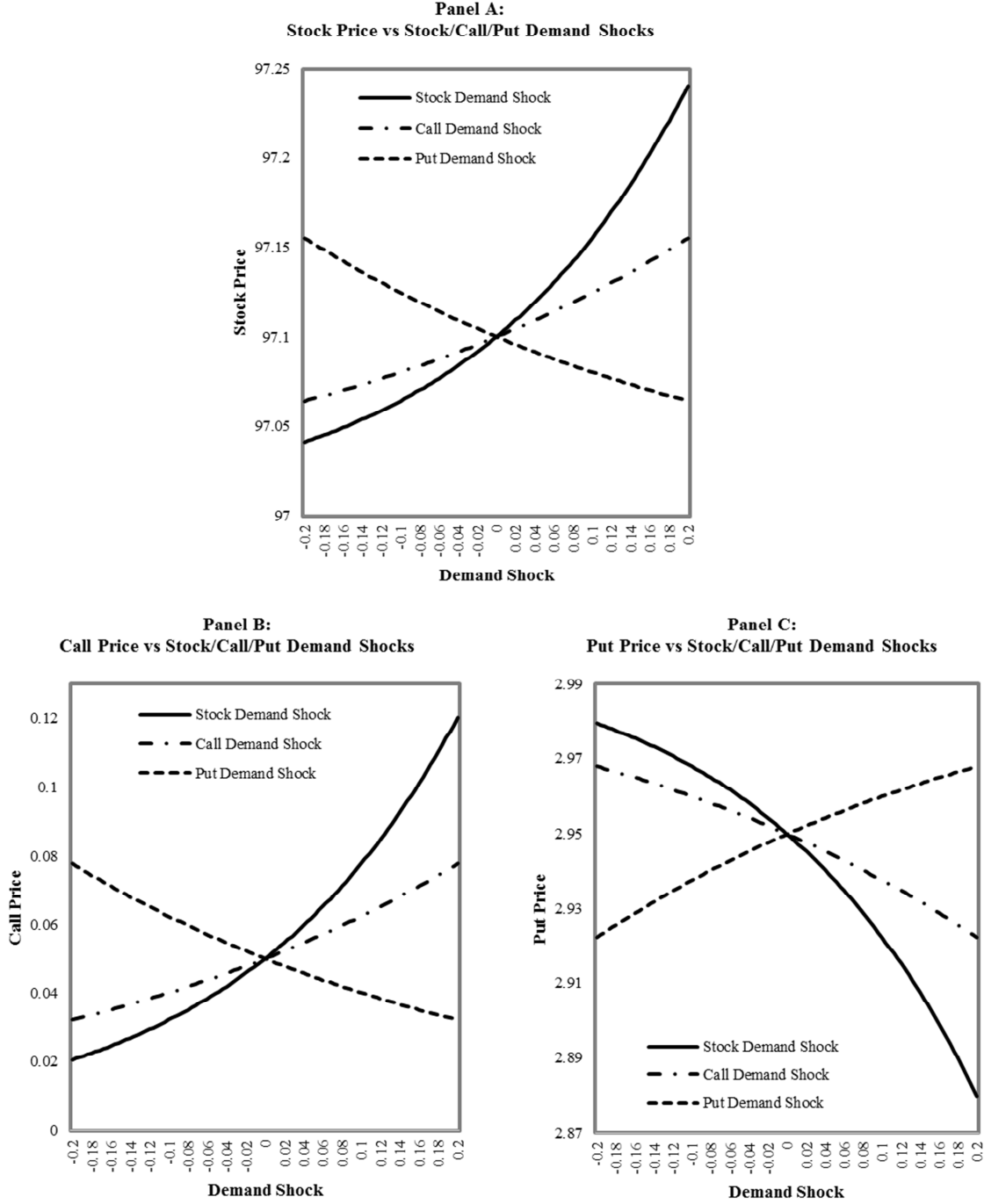


Figure 2: Stock and Options Prices and Noise Trading Shocks

This figure plots the prices of the stock, the call option, and the put option, as a function of uninformed demand shocks given a good signal, i.e., $\theta = 1$. Panel A plots the price of the stock as a solid line for stock demand shocks z , while keeping the call and put demand shocks at $\nu_c = 0$ and $\nu_p = 0$. The dash-dotted line represents the stock price as a function of call demand shocks ν_c , while holding the stock and put demand shocks at $z = 0$ and $\nu_p = 0$. The dotted line represents the stock price as a function of put demands shocks ν_p , holding the stock and call demand shocks at $z = 0$ and $\nu_c = 0$. Panels B and C repeat the exercise but plot instead the prices of the call and put options, respectively. The parameter values used to generate these plots are as follows: $F_H = 103$, $F_L = 97$, $K = 100$, $\omega = 0.7$, and $\gamma = 1.5$.

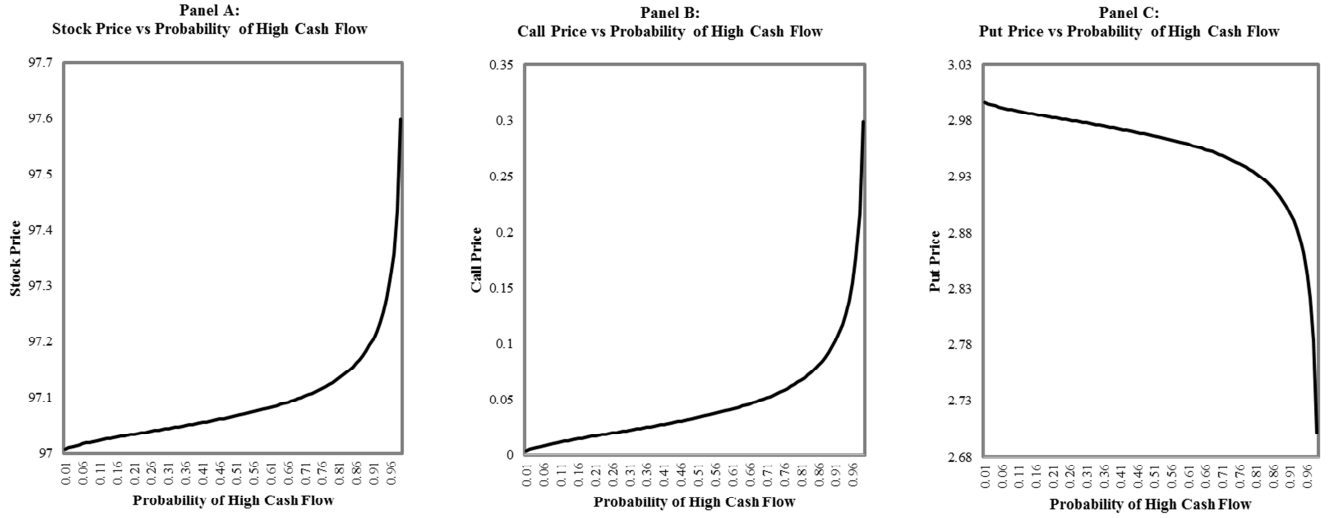


Figure 3: Comparative Statics on Probability of High Cash Flow

This figure shows the sensitivity of the stock price (Panel A), the call price (Panel B), and the put price (Panel C), to changes in the probability of the high cash flow (ω), given a positive signal ($\theta = 1$). Figure 2 shows that the curves intersect at the point in which all the noise trading shocks are equal to zero (i.e., $z = 0$, $\nu_c = 0$, and $\nu_p = 0$). This figure plots these intersection points for different values of ω . The parameter ω can be interpreted as the quality of the positive signal. Its baseline value is $\omega = 0.7$, and this figure takes it to vary in the interval $]0, 1[$. The remaining parameter values are kept unchanged: $F_H = 103$, $F_L = 97$, $K = 100$, and $\gamma = 1.5$.

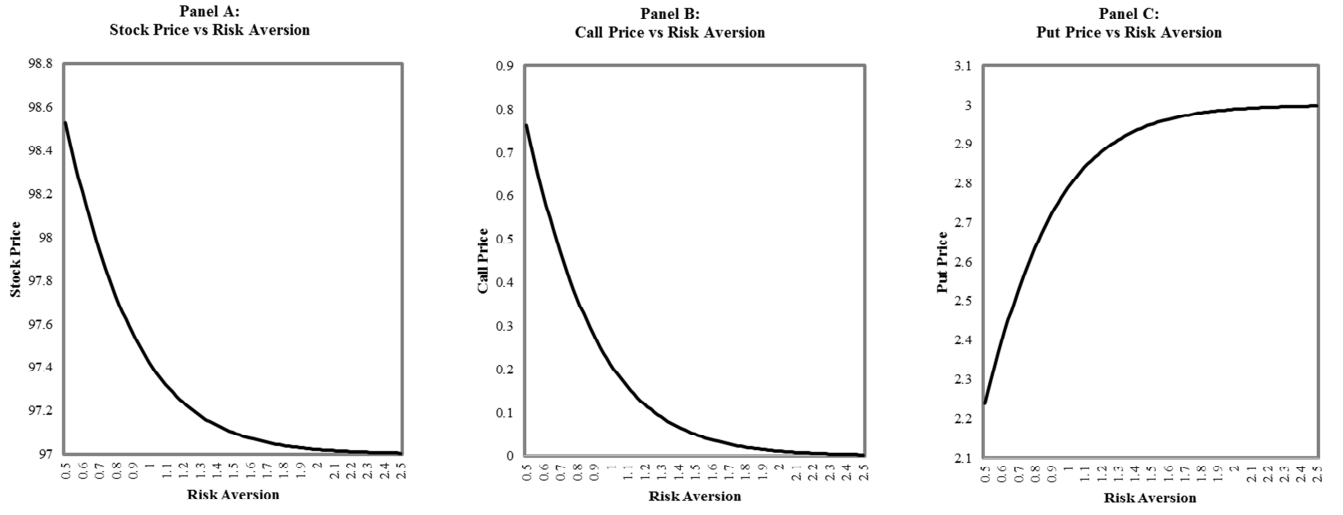


Figure 4: Comparative Statics on Risk Aversion

This figure shows the sensitivity of the stock price (Panel A), the call price (Panel B), and the put price (Panel C), to changes in the risk aversion parameter (γ), given a positive signal ($\theta = 1$). Figure 2 shows that the curves intersect at the point in which all the noise trading shocks are equal to zero (i.e., $z = 0$, $\nu_c = 0$, and $\nu_p = 0$). This figure plots these intersection points for different values of γ . Its baseline value is $\gamma = 1.5$. The remaining parameter values are kept unchanged: $F_H = 103$, $F_L = 97$, $K = 100$, and $\omega = 0.7$.

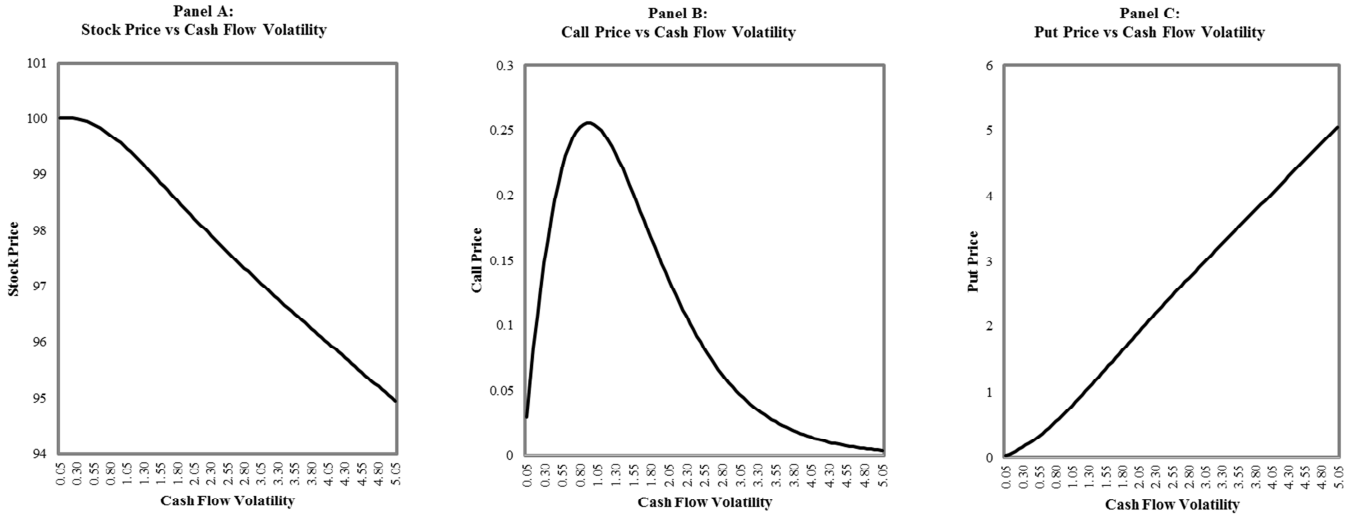


Figure 5: Comparative Statics on Cash Flow Volatility

This figure shows the sensitivity of the stock price (Panel A), the call price (Panel B), and the put price (Panel C), to changes in the volatility of the firm cash flows at time $t = 2$. The cash flow volatility is captured by the distance between the high and the low cash flows, i.e., $F_H - F_L$. The variation in cash flows is kept symmetric in relation to the strike price $K = 100$, which means that $F_H = K + a$ and $F_L = K - a$. In the baseline model, $F_H - F_L = 103 - 97 = 6$, which corresponds to the case $a = 3$. This figure plots the results for values of a between 0.05 and 5.05. Figure 2 shows that the curves intersect at the point in which all the noise trading shocks are equal to zero (i.e., $z = 0$, $\nu_c = 0$, and $\nu_p = 0$). This figure plots these intersection points for different values of a . The remaining parameter values are kept unchanged: $K = 100$, $\omega = 0.7$, and $\gamma = 1.5$.