Do Informed Investors Time the Horizon? Evidence from Equity Options

Selwyn Yuen*

This Version: April 7, 2014.

Abstract

I document evidence that equity options with different maturities do not embed the same type of stock information with the same strength. Long-dated options not only embed long-horizon stock information, but also *short*-horizon information, even if a cheaper and more liquid short-dated option is available. More surprisingly, options can even embed stock information *beyond* their expiration dates. I call these unique patterns *horizon timing effects*. These effects can arise in a stylized setting when informed investors optimally pick from a menu of option maturities in face of uncertainty in the timing or speed of price discovery, and the stylized predictions are matched closely with data. Specifically, I show that the differential information across options of different maturities can be operationally extracted and decomposed using option-based variables, revealing a clear and robust pattern of horizon timing effects. Validation tests are further performed to strengthen the information-based origin of these effects. The results are also robust to other confounding factors, such as microstructure effects and volatility effects. Overall, these findings suggest that informed investors form expectations not only on directional stock returns, but also on *when* returns will be realized, and accurately do so up to an impressive 6-month period in the future.

*Kellogg School of Management, Northwestern University: s-yuen@kellogg.northwestern.edu.

I am very grateful to helpful comments from Torben Andersen, Snehal Banerjee, Zhuo Chen, Jesse Davis, Zhenduo Du, Janice Eberly, Nicola Fusari, Naveen Gondhi, John Hull, Ravi Jagannathan, Raymond Kan, Robert Korajczyk, Arvind Krishnamurthy, Binying Liu, Andrea Lu, Brian Melzer, Tatjana Puhan, Paola Sapienza, Yao Shen, Brian Weller, Viktor Todorov. I also thank the Social Sciences and Humanities Research Council for financial support.

1 Introduction

The presence of informed trading in the option market has been widely documented, confirming that options can indeed embed stock return information.[1] However, does the strength or type of information vary across options of different maturities? If so, how should these differences in information sets across maturities be interpreted? This paper addresses these questions by providing evidence that options with different maturities do not embed the same type of information with the same strength, and suggests that these effects should be interpreted as informed investors trying to time the horizon. By *horizon timing*, I mean that investors form expectations of *when* returns will be realized. The implications of horizon timing are summarized by two predictions in my model. First, longer-dated options should predict not only long-horizon returns, but also short-horizon returns, even if a cheaper and more liquid short-dated option is available. Second, shorter-dated options should predominately predict short-horizon returns, but may predict returns beyond their expiration dates with a low but non-zero probability. Since these predictions arise from investors trying to time the horizon, I call these predictions *horizon timing effects*. The contribution of this paper is to test and confirm that the specific predictions of horizon timing effects are matched closely with data.

Unlike stock investors, informed investors in the option market are forced to take a view on when information will be reflected in stock returns by choosing among a menu of option contracts with different maturities. Being right in horizon timing is as important as being right in the *direction* of stock price movements, because otherwise, their options can expire worthless. In contrast, informed investors in stocks who is wrong about horizon timing can simply hold onto their position a little longer, as long as their directional prediction is correct. This unique property of the derivatives market provide a natural venue to identify horizon timing behavior of investors.

The option market can be preferred to the stock market for a few reasons, including the ability to

[1]E.g. Easley, O'Hara and Srinivas (1998), Lee and Yi (2001), Chakravarty, Gulen and Mayhew (2004), Cao, Chen and Griffen (2005), Pan and Poteshman (2006), Goyal and Saretto (2009), Cremers and Weinbaum (2010), Xing, Zhang and Zhao (2010), An et. al. (2013), Lin, Lu and Driessen (2013), etc.

short, to take leverage, or to hide trades when the option market is liquid. If investors with private information trade in options rather than stocks, information will be impounded into option premiums through trades and order flows in the option market *before* the stock market, assuming price discovery is not instantaneous. Easley, O'Hara and Srinivas (1998) build an asymmetric information model in which this lead-lag relationship between options and stocks results. If informed investors who believe that information will be impounded quickly into prices choose to trade shorter-dated options and vice versa, one would expect longer-dated options to embed longer-horizon information and shorter-dated options to embed shorter-horizon information. However, there are at least a few issues that complicate this seemingly intuitive story. First, shorter-dated options allow for a more levered position due to their cheaper premium and are more liquid with a lower transaction cost. Second, a conservative trader with private information may err on the side of choosing a longer-dated option, to hedge against the possibility that price discovery may be slower than they expect. These two opposing factors create a sweet spot for the optimal maturity. A third complication is that even if informed investors choose a shorter-dated option and "miss" (i.e. the private information has not yet been impounded into prices by the expiration date), they can still roll to the next option by paying additional transaction costs. A fourth complication can arise if an investor is highly confident that a piece of private information will not be leaked to the market before a certain date, she may well choose to wait till that date to invest in the then more liquid shorter-dated option. This would confound long-dated options' ability to predict long-dated returns. Section 2 discusses the economic intuition of these considerations, and Appendix B formalizes the intuition with a stylized model.

My baseline empirical results match the predictions closely, and also survive a battery of robustness tests and control variables. I operationally extract and decompose the differential information across options of different maturities by using an option-based predictor documented by Cremers and Weinbaum (2010). Their predictor, called put-call parity deviation (denoted $PCPD$), is computed as the difference between implied volatilities of the calls (CIV) and puts (PIV). They claim that investors with private positive information can either buy calls or sell puts, but both actions increase $PCPD$ as defined, and vice versa for investors with private negative information. They find that a higher $PCPD$ strongly predicts a cross-sectionally higher stock return, and show that this is due to information-based trading in the option market.

The horizon timing effects can be observed up to an impressive 6-month horizon, suggesting that part of the information gap is not closed quickly. This is not surprising if this information is private to informed investors. However, even if some of this information is public, it needs not contradict market efficiency in light of recent literature on slow-moving capital or optimal inattention with transaction cost [Abel, Eberly and Panageas (2007, 2013), Duffie (2010), Alvarez, Guiso and Lippi (2013)]. On the other hand, delayed response to information that persists for a few months has been documented, such as the post-earnings announcement drift [Bernard and Thomas (1989, 1990), Jegadeesh and Livnat (2006)].

The information-based origin of horizon timing effects can be validated by certain conditioning variable tests. Roll, Schwartz and Subrahmanyam (2010) investigate the option-to-stock volume ratio (O/S), and find that when O/S is high, a stronger presence of information should be found in options. I build on these prior works to show that horizon timing effects also strengthen for higher O/S, a trend consistent with information-based interpretations. Similarly, I condition on an option turnover measure ($O/SHROUT$), computed as option volume divided by number of shares outstanding. Intuitively, it captures how fast the option market can turnover the inventory of stocks in the market, which is an alternate proxy for option liquidity. I find that the horizon timing effects strengthen with option turnover, as expected.

The rest of this paper is organized as follows. Section 2 theoretically motivates the horizon timing effects and forms testable predictions. Section 3 describes the data. Section 4 presents my baseline empirical results with validation tests. Section 5 confirms the robustness of these observations. Section 6 explicitly rejects alternative explanations. Section 7 concludes.

2 Motivation

The economic intuition underlying the horizon timing effects can be explained in a simple setting. These intuitions are formalized into a simple model in Appendix B. The intent here is *not* to motivate why options embed stock information, which has already been done,[2] but rather, to motivate why there could be *differential* information across options of different maturities on various horizon returns. I assume that an informed investor has private information regarding future stock returns, but is uncertain about when returns will be realized. This uncertainty comes from at least two different sources: (1) uncertainty in *news-release timing*, and (2) uncertainty in *price-discovery timing or speed*. Therefore, even if a news-release date is known in advance, there is still uncertainty in how long it takes for information to be impounded into prices. Although I am not aware of any existing models that focus on a stochastic speed or timing of price discovery, the assumption does not seem hard to justify. Previous studies claim that the speed of price adjustment to news is a function of size, analyst coverage, institutional ownership, trading volume, relative order of news announcement, and foreign ownership [Brennan et al. (1993), Badrinath et al. (1995), Chordia and Swaminathan (2000), Andersen et. al. (2003), Hou (2007), Cai et. al. (2013)]. Many of these variables can be stochastic. Andersen et. al. (2007) even find evidence that the stock market reacts to the same news differently depending on the state of the economy. All these previous works support a stochastic price-discovery time and speed.

Consider a 2-period setting with a short horizon T_1 and a long horizon T_2, with $0 < T_1 < T_2$. Suppose a risk-neutral informed investor has private information that a stock will payoff $R > 0$ within the period $t \in [0, T_2]$, but is otherwise uncertain about whether the payoff will arrive in the short or long-horizon. In particular, the agent knows at time $t = 0$ that the stock will either payoff R in period $t \in [0, T_1]$ with probability p or period $t \in (T_1, T_2]$ with probability $1 - p$. If the payoff R comes in the first period, the payoff in the second period is zero, and vice versa. Risk-free rate is normalized to zero to eliminate discounting across the two periods. At time $t = 0$, the agent is

[2] Easley, O'Hara and Srinivas (1998) provides a classic model to explain the lead-lag effects between option and stock market. An et. al. (2013) extend their model to explain similar effects in their first-difference IV-based predictors.

endowed with wealth W, and makes a decision to either: (1) invest in long-dated at-the-money (ATM) call options, (2) invest in short-dated ATM call options, or (3) not invest in either.[3] At time $t = T_1$, the agent re-decides again whether she will continue to hold the option (if she invested in the long-dated option), roll her option to the next short-dated option (if she invested in the short-dated option), or invest in the short-dated option (if she did not invest before).

According to bid-ask spread statistics (to be presented in Section 3.2), transaction cost of options increases in maturity. Without loss of generality, I assume that transaction cost is not incurred when an option expires naturally. Let Δ_1 and Δ_2 denote the deltas, c_1 and c_2 denote the premiums, and F_1 and F_2 denote the one-way transaction costs of short and long-dated options respectively, therefore $c_2 > c_1$ and $F_2 > F_1$. Using a first-order approximation, option value will change by ΔR when stock price changes by R. In this simplified setting, the core intuition can be summarized as follows. When p is small ($p < P_{low}$), she prefers to wait till the second period to re-decide. When p is non-extreme ($p \in [P_{low}, P_{high}]$), she wants to buy the long-dated option. When the probability of payoff in the first period p is high ($p > P_{high}$), the informed investor wants to buy the short-dated option. Appendix B formalizes these arguments and show that under certain parameter conditions, $0 < P_{low} < P_{high} < 1$, and thus, three regions of investment decisions are created. These observations are captured in two stylized facts below.

Stylized Fact 1: *Informed investors will trade long-dated options specifically when the timing of payoff is uncertain (i.e., when p is not extreme), provided payoff R is sufficiently large. Therefore, long-dated options should embed stock information of both periods.*

The fact that longer-dated options embed second-period information may be intuitively obvious. What is less obvious is that it also contains short-horizon information, *even if a cheaper and more liquid short-dated option is available*. In fact, the content of Stylized Fact 1 is that long-dated options must contain *both* short and long-horizon information, not just solely long-horizon information. Further, short-dated options can be chosen even when the informed investor believes that there is

[3] Call options are used for explanation here, but the argument works for either buying calls or selling puts when $R > 0$, and buying puts or selling calls when $R < 0$.

some probability (i.e., $1-p>0$) that payoff will occur in the second period. This is captured in the second stylized fact.

Stylized Fact 2: *Short-dated options contain predominantly short-horizon stock information, but may, with a low but non-zero probability, contain long-horizon information. In other words, short-dated options investor can be "wrong" in the sense that payoff may occur beyond the chosen maturity.*

Again, the fact that short-dated options embed first-period information is intuitive. The more interesting prediction of Stylized Fact 2 is that short-dated options can also embed a small amount of second-period information. Taken together, these two stylized facts provide an operational definition of *horizon timing effects*—the central claim of interest in this paper—and their specific predictions can be tested using option and stock data. These predictions form a "lower diagonal" pattern, as summarized in this table.

Horizon Timing Effects in a 2-period Model

	Short-horizon Return	**Long-horizon Return**
Short-dated Options	Strong	Weak, but non-zero
Long-dated Options	Strong	Strong

Define $\gamma \equiv \frac{c_1+F_1}{c_2+F_2}$ to be the ratio of the costs of acquiring the short to long-dated options assuming that we buy the options.[4] The ratio γ is interesting here because it is the *relative* leverage across the short and long-dated options. For example, if the long-dated option is 30% *less* levered than the short-dated option, but the probability of the second-period payoff is 40%, then the long-dated option is still preferred since the forgone second-period payoff is more penalizing to the short-dated option than the extra leverage achieved. This intuition, however, does not take into account of the fact that a long-dated option can be sold with a residual time value at $t = T_1$. Taking the residual time value of the long-dated option into account will result in a slightly higher indifference probability than γ, i.e. $P_{high} \gtrsim \gamma$.

[4]This assumes that we buy the options at the ask price. The argument also works if we sell instead of buy the options, but γ would be defined as $\frac{c_1-F_1}{c_2-F_2}$ in that case.

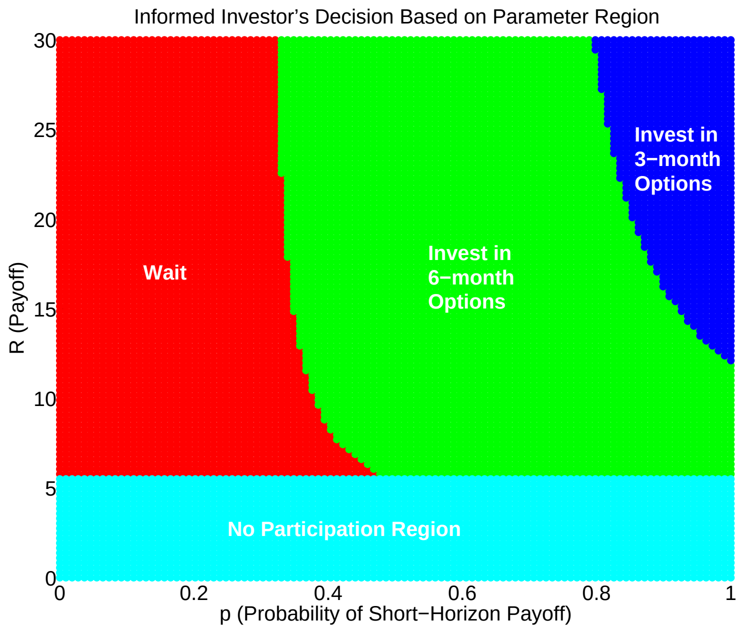
2.1 Calibration

I next calibrate these parameters with some empirical data. In Section 4.2, I use 3 and 6-month options to empirically test these horizon timing effects, so $T_1 = 3$ and $T_2 = 6$ months respectively. The goal is to calibrate the parameters F_1, F_2, c_1 and γ, and express the optimal investment decision based on parameter region in the (p, R) space. Based on Panel C of Table 2 (detailed in Section 3), the median bid-ask spread of options is roughly $\$0.37$. I take this as an approximation of 3-month spreads. I calibrate F_2/F_1 from Figure 1, which shows that the 6-month bid-ask spread is on average 19% bigger than F_1. Under this calibration, $F_2/F_1 = 1.19$ and the *one-way* transaction cost F_1 is half of bid-ask spread, which is $F_1 \approx 0.37/2 = 0.185$, and $F_2 \approx F_1 \times 1.19 = \$0.22$. Based on the standard Black-Scholes formula, $\gamma \equiv \frac{c_1+F_1}{c_2+F_2} \approx \frac{c_1}{c_2} \approx 0.7$ (stock price $S = \$50$, strike $K = Se^{(r+\sigma^2/2)T}$ for ATM options explained in previous footnote, interest rate $r = 3\%$, and volatility $\sigma = 30\%$), and is in fact quite insensitive to the exact parameters used. These parameters also imply that $c_1 \approx \$2.7$. The figure below (on next page) illustrates how the optimal investment decision changes across the parameter regions, for different p and R. The leftmost region represents the decision to wait till $t = T_1$, the middle region represents the decision to invest in the long-dated option, and the rightmost region represents the decision to invest in the short-dated option.

3 Data

Stock price and volume data are from CRSP, stock fundamental data come from Compustat, and options data are from OptionMetrics. OptionMetrics provides implied volatility surfaces of individual equity options based on standardized expirations that correspond to 1-6, 9, 12, 18, and 24 months and standardized moneyness that correspond to deltas of 0.20, 0.25, 0.30, 0.35, ..., 0.80. Therefore, ATM options are defined by an (absolute) delta of 0.5[5] and OTM options are

[5] From standard Black-Scholes formula and notations, this implies that $\Delta \equiv \frac{\partial C}{\partial S} = N(d_1) = 0.5$, or $d_1 = 0$, where $d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right]$. In other words, at-the-money call options have strike $K = Se^{\left(r+\frac{\sigma^2}{2}\right)T}$, and similarly for puts.



defined by an (absolute) delta of 0.2. Even though the volatility surface goes up to 24 months, I will show in Section 3.2 that liquidity is extremely limited for options beyond eight months. As a result, I will only use options up to 6 months to expiration in this study. Implied volatilities are provided for calls and puts separately. The OptionMetrics volatility surface is computed in two steps. First, implied volatilities are computed using a classic binomial tree method [Cox, Ross and Rubinstein (1979)], taking into account of dividends and early exercise premiums of each equity option, which are American options. For each strike and maturity, the mid price (average of last bid and ask of the day) is used to compute the implied volatilities for each option contract traded. Second, to obtain the implied volatility surface on the standardized grid, kernel smoothing is performed at each standardized point based on a weighted sum of all points that are in proximity, where proximity is defined in terms of (1) distance in log maturity, (2) distance in deltas, and (3) an indicator function of whether call-put type matches. Further, a standardized option is included only if a sufficient number of option data points are available for interpolation. The volatility surface data is particularly convenient for this analysis because it contains implied volatility for a fixed set of deltas and maturities, obviating the need to standardize for proper comparisons across individual options. The volatility surface data provided is *complete* in the sense that once an individual firm is selected for a particular date, implied volatility data of *all* standardized deltas and expirations are guaranteed to be non-missing for that firm-date.

Option volume and open interest data are obtained from the daily Option Price tables in OptionMetrics. Daily option volume and open interests are aggregated over each month by summing daily volumes and open interests to obtain monthly option volume and open interests respectively. Monthly stock returns, stock volume, and market value of equity are from standard CRSP files. Since the options data have the shortest history among the three data sources, starting in 1996, the merged data cover the period of 1996 to 2011. The last year of data in 2012 is excluded because N-month forward returns ($N = 1, 2, \ldots, 12$) are needed in my computations defined later. To construct the universe of securities, I include only common stocks from NYSE, NYSE MKT (formerly AMEX) and NASDAQ in CRSP that are matched to an option in OptionMetrics. ADRs, ETFs, and closed-end funds are all excluded. The merged dataset has $273,487$ observations, with the number of securities averaging around $1,700$ per year. Table 1 shows the number of securities in the merged dataset by year.

3.1 Put-call Parity and $PCPD$

Put-call parity [Stoll (1969)] is a no-arbitrage relationship for European options that requires minimum asset-pricing assumptions to hold. It states that a long position in a European call and a short position in a European put of identical strike can be replicated by holding a unit of the underlying stock and borrowing the present value of the strike. However, call and put prices may deviate from this relationship for at least a number of reasons, including dividends (Div) and early exercise premium (EEP). Cremers and Weinbaum (2010) point out that the following relationship is no longer a no-arbitrage relationship because of the explicit incorporation of the early exercise premium in the equation.

$$C - P = S - PV(K) - PV(Div) - EEP(P) + EEP(C) \tag{1}$$

Since individual equity options are American options, the implied volatilities of calls and puts that take into acount of dividends and early exercise premiums should be identical. Other market imperfections can further affect the validity of this relationship, including bid-ask spread, price impact, tax, availability of stock borrows, difference between borrowing and lending rates, margin requirements, etc. Therefore, the put-call parity as in (1) is not a strict equality in reality, but rather, has some "thickness", within which no arbitrage opportunities exist. Therefore, deviations from the put-call parity relation do not imply that put-call parity is being violated. Assuming that informed investors trade in the option market, bullish investors will buy calls or sell puts, both actions widening the gap between call and put implied volatilities (CIV and PIV). I define put-call parity deviations ($PCPD$) as the difference between the implied volatilities of at-the-money (ATM) call and put. The subscript "ATM" is dropped whenever there is no ambiguity in the context. Option predictors based on the last trading day of each month are extracted for the purpose of computing a predictive measure for next month's stock returns.

$$PCPD_{ATM} \equiv CIV_{ATM} - PIV_{ATM} \tag{2}$$

Other studies also propose other IV-based predictors of stock returns. Xing, Zhang and Zhao (2010) define "smirk" as the IV of out-of-the-money put minus that of ATM call, and show that this measure predicts significantly negative stock returns. They argue that bearish investors not only tend to buy puts, but especially OTM puts, a strategy that is cheaper to implement and allows for a higher leverage. Their skew measure can be reinterpreted as a $PCPD$ and a slope component:

$$\begin{aligned} SMIRK &\equiv PIV_{OTM} - CIV_{ATM} \\ &= \underbrace{(PIV_{OTM} - PIV_{ATM})}_{PSKEW} - \underbrace{(CIV_{ATM} - PIV_{ATM})}_{PCPD_{ATM}} \end{aligned}$$

Figure 2 depicts these IV-based measures qualitatively in an IV smile diagram. In Appendix C, I provide empirical evidence that the predictive power of $SMIRK$ comes from $PCPD$ but not from $PSKEW$. This is consistent with Conrad, Dittmar and Ghysels (2013), who also finds $SMIRK$

to be a noisy measure. An et. al. (2013) further investigate the predictive power of the change in CIV, PIV, and $PCPD$ over time (i.e. ΔCIV, ΔPIV and $\Delta PCPD$) and conclude significant predictive power in these first-differences. Their interpretation is that these first-differences proxy news arrival. $PCPD$ is suitable for use in this paper instead of other IV-based measures because it is the most robust and persistent predictor of the three, as demonstrated in Appendix C.

3.2 Summary Statistics

Figure 1 presents option volume distribution and bid-ask spreads across maturities. The top figure shows a distribution of option volume across maturities, and the bottom figure shows the ratio of median bid-ask spreads between the Nth month and the first month.[6] I confirm the well-known fact that longer dated options have lower volume (liquidity) and higher transaction cost on average. In fact, options beyond eight months maturity are so thinly traded that information is not expected to be present for those maturities. For this reason, I use only options up to 6 months to maturity in this study.

Panel A of Table 2 summarizes the mean, standard deviation, skewness and correlation matrix for both implied volatilities (IV) of ATM calls and puts (CIV, PIV) of three different maturities: 1-month, 3-month, and 6-month. Since this is a panel dataset, all summary statistics are first computed for individual cross-sections, and then averaged over time. The correlation of CIV and PIV is quite high at 0.91, giving some support that put-call parity is at least approximately true. This is consistent with the results reported in An et. al. (2013). The IV's across different maturities are also highly correlated. However, Panel B of Table 2 verifies that the correlations of $PCPD$'s across different maturities are less correlated than before, at 0.81. On average, $PCPD$ has a negative mean and skewness, meaning that the IV of puts are generally higher. The mean absolute deviation (MAD) is also displayed in the bottom row, indicating that the IV's of calls and puts deviate from each other by about 2-3% on average. It is this deviation that embeds

[6]The volume distribution and median bid-ask spread ratio are first calculated for each firm-date, then averaged over all firm-dates. There are alternate ways to summarize the data, but the qualitative results do not change.

stock return information. $PCPD_{6\perp 3}$ is the orthogonalization of $PCPD_6$ with respect to $PCPD_3$, computed as the residuals of regressing $PCPD_6$ on $PCPD_3$ with an intercept, resulting in a (manufactured) correlation of zero between $PCPD_{6\perp 3}$ and $PCPD_3$. Despite the correlation of 0.81 between $PCPD_3$ and $PCPD_6$, I note that the residual standard deviation in $PCPD_{6\perp 3}$ is still about half of $PCPD_6$, meaning that residual variability remains to explain future stock returns. Section 4 gives a more detailed interpretation on this orthogonalization procedure.

Panel C of Table 2 reports the summary statistics of bid-ask spreads in stocks and options. The stock's bid-ask spread is defined as $SBAS = \text{Ask} - \text{Bid}$ with a median of $\$0.13$, and proportional bid-ask spread is defined as $\%SBAS = SBAS/\left(\frac{\text{Ask}-\text{Bid}}{2}\right)$ with a median of 51 basis points. The option market is more expensive to trade, with $OBAS$ having median around $\$0.37$. Not surprisingly, all spreads are right-skewed, indicating the presence of abnormally high spreads for some very illiquid stocks and options.

4 Empirical Results

4.1 Methodology

I first clarify some relevant notations. Appendix A contains a more complete list of notations. R_N denotes the Nth-month percentage stock return, R_{1-N} denotes the cumulative N-month percentage return divided by the number of months N, and R_{N-M} denotes the Nth-to-Mth month percentage stock return, divided by the number of months $N-M+1$. The subscript of $PCPD_{N,ATM}$ means that this $PCPD$ measure is computed based on at-the-money N-month options. The subscript "ATM" is sometimes suppressed. Finally, predictors computed based on long-dated and short-dated options are called "long-dated predictors" and "short-dated predictors" respectively.

Fama-MacBeth regressions is performed on the panel data, with all standard errors and t-statistics being Newey-West adjusted with 6 lags in this paper [Newey and West (1987)]. In the first stage, coefficients are obtained by running cross-sectional regressions of next-period stock return on the option-based predictor $PCPD$ with controls, as follows:

$$R_{1-N,i} = \delta_t + \beta_t \cdot PCPD_{i,t} + \lambda_t \cdot \text{Controls}_{i,t} + \varepsilon_{i,t} \tag{3}$$

These coefficients are then averaged over time to obtain the Fama-MacBeth coefficients. The subscripts i and t indexes stock and month respectively. Similar to other studies in this literature [An et. al. (2013), Xing, Zhang and Zhao (2010)], I control for some common asset-pricing factors. For each month, $BETA$ is computed by regressing daily excess returns of individual stocks on the daily excess returns of the market. $SIZE$ is the log of market capitalization. $B2M$ is the log of book-to-market ratio. MOM is past 2-to-12 month cumulative return, i.e. momentum. $ILLIQ$ is Amihud's illiquidity, computed by dividing absolute monthly stock return by monthly stock volume. Short REV is short-term reversal, proxied by past 1-month return. $RVOL$ is the standard deviation of daily returns within each month. $PSKEW$ is the left-skew based on puts discussed in Section 3.1, computed as $PIV_{OTM} - PIV_{ATM}$. V_c/V_p and OI_c/OI_p are the call/put volume and open-interest ratios respectively. These are my standard controls.

My baseline results will be based on the following Fama-MacBeth regression, where the dependent variable is now $R_{N-M,i}$, as defined before.

$$R_{N-M,i} = \delta_t + \beta_t \cdot (PCPD_3)_{i,t} + \gamma_t \cdot (PCPD_{6\perp 3})_{i,t} + \lambda_t \cdot \text{Controls}_{i,t} + \varepsilon_{i,t} \tag{4}$$

$PCPD_3$ is the 3-month predictor. $PCPD_{6\perp 3}$ represents $PCPD_6$ orthogonalized to $PCPD_3$, computed as the residual of regressing $PCPD_6$ on $PCPD_3$ with an intercept. Interpreting $PCPD$ as a signal, this orthogonalization gives the marginal information of one signal net of another. The two stages of Fama-Macbeth regressions imply that this orthogonalization procedure is performed at each cross-section independently.[7]

[7] In Section 5, I confirm that my results are robust to various orthogonalization methods, and that my conclusions are not driven by the specifics of how I orthogonalize these signals.

4.2 Baseline Results

In column 1 of Table 3, I regress R_1 on $PCPD_{1,ATM}$ without any controls. The coefficient of $PCPD_{1,ATM}$ is 5.90 with a t-statistic of 7.01, which is significant at the 1% level, confirming the predictive power of $PCPD$ in Cremer and Weinbaum (2010). In column 2, this predictive power continues to hold with controls. Columns 3-4 confirm analogous results by regressing R_{1-3} on $PCPD_3$. All results are significant at 1% level. None of these is surprising given what has been documented previously. The coefficient of Short REV is negative and significant in column 2, but this significance disappears in column 4, indicating that short-term reversal is a short-lived effect that quickly dissipates beyond 1 month. The coefficient of realized volatility ($RVOL$) is negative and marginally significant in columns 2 and 4, which is consistent with the cross-sectional volatility effects found by Ang et. al. (2006, 2009), although they focus on idiosyncratic volatility.

Table 4 serves as the baseline result of horizon timing effect. In columns 1-2 of Panel A, I regress R_{1-3} and R_{4-6} on both $PCPD_3$ and $PCPD_{6\perp3}$ without any controls. As discussed before, this decomposition tests for any *extra* predictive power coming from $PCPD_6$ on top of $PCPD_3$. $PCPD_3$ in this case can be interpreted as the base, and its coefficient value of 4.69 is unchanged from Table 3.[8]. This empirical result matches closely the stylized facts discussed in Section 2 First, in columns 1-2, the coefficients of $PCPD_{6\perp3}$ are 6.14 and 6.16 respectively. These coefficients have comparable magnitudes and significance, which imply that $PCPD_{6\perp3}$ embed information in *both* the short and long horizons. This is precisely Stylized Fact 1. It also suggests that informed investors take into consideration of *horizon timing* in their options trading activities. Indeed, the predictive power of longer-dated options is so impressively persistent that it can predict return 4 to 6 months in the future. Second, the coefficients of $PCPD_3$ are 4.69 and 0.99 respectively, implying that the predictive power of the $PCPC_3$ is dramatically weaker beyond the 3-month maturity date. This matches closely with Stylized Fact 2, with a distinct lower-diagonal pattern in the coefficients. Columns 3-4 confirm similar facts and pattern by including controls. The weaker but significant coefficient (1.32) in the upper-right corner does not contradict my predic-

[8]The coefficient of $PCPD_3$ in column 3 of Table 4 differs slightly from that of column 4 of Table 3 because the orthogonalization procedure is not done with controls.

tions, but actually attests to the fact that shorter-dated options can embed information beyond its maturity with a *small* probability. Since the orthogonalization procedure can shrink the variance of $PCPD_{6\perp 3}$, as discussed in Section 3.2, one may worry that this result is biased by scaling differences between the two predictors. To confirm that this is not the case, I hereafter normalize each regressor to a unit standard deviation and zero mean to ensure comparable scaling across $PCPD_3$ and $PCPD_{6\perp 3}$. The results in columns 5-6 shows that the magnitude of the coefficients changes due to the normalization as expected, but $PCPD_{6\perp 3}$ can still predict both R_{1-3} and R_{4-6} with significance at comparable magnitudes. The lower-diagonal structure is once again clearly observed. Finally, Panel B of Table 4 shows the regression of individual Nth-month return (R_N for $N = 1, 2, \cdots, 6$) on the predictors. It is clear that the predictability of $PCPD_3$ decays dramatically from 0.47 to 0.03, whereas the predictability of $PCPD_{6\perp 3}$ stays strong at around 0.12-0.13 throughout the 6 months. The *statistical* significance of $PCPD_3$ also decays quickly compared to $PCPD_{6\perp 3}$.

A more refined decomposition of this horizon timing effect is also possible using 2/4/6-month options. Table 5 parallels the previous result using $PCPD_2$, $PCPD_{4\perp 2}$, and $PCPD_{6\perp 4,2}$. $PCPD_{6\perp 4,2}$ is computed by taking the residuals of regressing $PCPD_6$ on both $PCPD_4$ and $PCPD_2$ with an intercept. Columns 1-3 of Panel A display results without controls, columns 4-6 include the standard controls, and columns 7-9 normalize the regressors as before. Once again, the 3-by-3 matrix of coefficients has significant lower-diagonal elements (Stylized Fact 1). The upper diagonal coefficients are occasionally statistically significant, but much smaller in magnitudes (Stylized Fact 2). Panel B of Table 5 also parallels that of Table 4 with R_N as the dependent variable, and a differential persistence can be similarly identified as before. However, there is a limit to how refined this decomposition can go. For example, I find that a decomposition with 1/2/3/4/5/6-month options may be overly demanding for the current regression specification, and the results become too noisy to yield observable pattern. This may suggest that informed investors can only time the horizon up to an accuracy of a 2-month wide window, but not more refined.

Figure 3 presents a compelling plot that shows predictors of different maturity have different speed of decay of predictive power. Each line in the figure represents using a separate predictor $PCPD_T$ as a standalone regressor. Each point on a line is a separate regression of R_{1-N} on each $PCPD_T$. The x-axis is N-month return, and the y-axis is the Fama-MacBeth regression coefficient of the respective $PCPD_T$. From this figure, shorter-dated predictors clearly decay faster than longer-dated predictors.

4.3 Validations

4.3.1 Validation Test #1: Effects of O/S Ratio

The results so far document the existence of horizon timing effects, but it is not clear that these effects necessarily come from informed trading. This section presents two additional evidences to strengthen the information-based origin of horizon timing effects. In their seminal paper, Easley, O'Hara and Srinivas (1998) present a model to explain how lead-lag effects can arise between the option and stock market. One intuition captured in their model is that informed traders prefer to trade in the option market when the option market is relatively liquid, the stock market is relatively illiquid, and the presence of informed traders is high. As a result, if the horizon timing effects are driven by informed trading activities, one would expect that the strength of the effects correlate negatively with stock volume and positively with option volume. However, volume measures are frequently too noisy when used alone. The option-to-stock volume ratio (O/S) measure reduces the noise by canceling out shocks that occur to both markets. Roll, Schwartz and Subrahmanyam (2010) find that the O/S ratio spikes before earnings announcement, and varies with analyst coverage or institutional ownership, suggesting that it is a potential measure of the informed trading in the option market.[9] My following validation test is based on these insights. I compute O/S using the aggregate monthly option-to-stock volume, and explore whether horizon timing effects become stronger for higher values of O/S. In the Fama-MacBeth regression, I separate each cross-section into three terciles sorting on O/S, and report the

[9]Johnson and So (2012) also investigate the O/S measure, and find that it not only reflect the presence of private information, but also predicts significant negative stock return. They relate O/S to short-sale costs.

Fama-MacBeth coefficients by tercile. Table 6 presents these results. Control variables have been excluded for these tests to avoid potential trending or interaction between the O/S and any of the control variates. In columns 1-3, R_{1-3} is the dependent variable, and the three terciles of O/S increase from left to right. Column 4 shows the difference between the top and bottom tercile, or (3)-(1). The interpretation of columns 5-7 and 8 are analogous, but for the R_{4-6} horizon. All coefficients trend upwards from left to right in columns 1-3 and 5-7, and the difference column is significant in column 8 at 10% level. Overall, these results strengthen the information-based origin of horizon timing.

4.3.2 Validation Test #2: Effects of Turnover

Another conditioning variable is option turnover ($O/SHROUT$), defined as the ratio of option volume to the number of shares outstanding in the stock. It captures the liquidity or tradability of options for each stock. The normalization by number of shares outstanding allows for a more comparable liquidity measure cross-sectionally and adjusts for some of the noise in volume. Although I am not aware of option turnover being used as a conditioning variable in this literature, stock turnover ($S/SHROUT$) has been used. In particular, Xing, Zhang and Zhao (2010) as an interaction variable with their IV-based predictor, with a hope to find that a *lower* stock liquidity implies stronger predictability. However, their results struggle with statistical significance. One possible problem with $S/SHROUT$ is that stock volume can be ambiguous in identifying the presence of informed trading in options. On the one hand, a lower stock volume could drive informed traders away from stocks to options, thus a stronger effect. On the other hand, the low stock volume can also be explained by a small presence of informed traders in that asset in general for both markets. This ambiguity makes $S/SHROUT$ ill-suited as a conditioning variable, but this ambiguity does not apply for $O/SHROUT$. Table 7 confirms this argument. Panel A of Table 7 parallels Table 6, but the conditioning variable is now $O/SHROUT$. Once again, the coefficients are clearly trending upwards from left to right for both columns 1-3 and 5-7, suggesting that a more liquid option market strengthens the horizon timing effect. In Panel B of Table 7, I check and confirm that $S/SHROUT$ is indeed ambiguous, by noting that no identifiable trends exist from left to right. Another reason of showing this invalid sort on $S/SHROUT$ is to demonstrate that the previous effective sorts on $O/SHROUT$ or O/S were not just fortunate coincidences. These evidences once again support the information-based origin of horizon timing effects.

5 Robustness

Table 8 presents various robustness tests. In columns 1-2, I run the Fama-Macbeth regression with normalized coefficients, except that extreme observations of CIV and PIV are trimmed at the 1% and 99% levels. Stylized Facts 1 and 2 remain, indicating that the current results are not driven by some outliers.[10] In columns 3-4, instead of excluding low-priced stocks below $5, all stocks are included. The results survive, indicating that horizon timing effects are not due to price filters. In columns 5-6 and 7-8, the sample period is divided into two equal sub-periods from 1996-2003 and 2004-2011. The results indicate that the predictability of $PCPD$ is stronger in the earlier sub-period, and weaker in the latter, although still significant in both periods. In contrast, Cremers and Weinbaum (2010) find that the significance disappears in the later years, and explains that this is due to the mispricing being gradually eliminated by market participants. I agree with their general interpretation, but note that the significance of the latter sub-period in my results differ from theirs because I specifically separate out the $PCPD$ of longer-dated options (3 and 6-month), whereas they perform a weighted average of put-call parity deviations across all strikes and maturities. My results indicate that for longer-dated predictors, this mispricing is not completely eroded even 6 years after their sample. Finally, different orthogonalization procedures are also tested. In columns 9-10, a simple subtraction $PCPD_6 - PCPD_3$ is used to replace the current orthogonalization. In columns 11-12, orthogonalization of $PCPD_6$ is done with respect to $PCPD_3$ and all standard controls ($PCPD_{6\perp 3,controls}$). Horizon timing effects are robust to both changes, indicating that they are not sensitive to the orthogonalization procedure used.

[10] It is important to remember that $PCPD_3$ is the base of decomposition while $PCPD_{6\perp 3}$ represents the *marginal* information on top of that. Thus, even if the magnitude of $PCPD_3$ (0.07) is getting closer to $PCPD_{6\perp 3}$ (0.09) in the R_{4-6} regression (column 2), the long-dated predictor is still significantly stronger than the short-dated predictor.

5.1 Portfolio Sorts

My findings can be verified with portfolio sorting techniques. Portfolio sorting is useful for several reasons. First, portfolio alphas can give a more direct measurement of *economic* significance. Second, it is another way to verify that the previous regression-based results are not driven by outliers. Third, it offers an extra robustness check. In particular, firms in the top decile of $PCPD$ should outperform the bottom decile of $PCPD$, after controlling for the Fama-French three factors and momentum.[11] In column 1 of Table 9, I confirm this by reconstructing the portfolio every month by sorting on $PCPD_1$. All numbers reported in the table are annualized portfolio alphas in percentage after controlling for the Fama-French three factors and momentum. The last row presents the portfolio alphas for decile 10 (high $PCPD$) minus decile 1 (low $PCPD$), which yields a large and significant annualized alpha of 21.78%,[12] again confirming its predicative power.

The portfolio sorting approach can also be used to test for horizon timing effects. To test for predictability in the R_{4-6} horizon, I skip $S = 3$ months, then hold the portfolio for another $H = 3$ months, rebalancing only $1/H$ of the portfolio every month. This portfolio construction methodology is similar to momentum [Jegadeesh and Titman (1993)]. In particular, Chen and Lu (2013) investigate how information in the option market can strengthen momentum returns. However, my result is different from momentum because I am sorting on $PCPD$, as opposed to past stock returns. In this section, I merely borrow the portfolio construction methodology from momentum to verify horizon timing effects. Columns 2-3 of Table 9 sort on $PCPD_3$ and show the annualized 4-factor alphas for both horizons, whereas the next two columns sort on $PCPD_{6\perp 3}$. Not surprisingly, sorting on $PCPD_3$ gives a large outperformance of 11.7% in R_{1-3}, but a much smaller outperformance of 3.6% in R_{4-6} (Stylized Fact 2). Sorting on $PCPD_{6\perp 3}$ generates an outperformance of 6.4% in the first 3 months, and 4.2% in the next 3 months (Stylized Fact 1). The outperformance of $PCPD_3$ in R_{4-6} (3.6%) is also the weakest of all (lower-diagonal pattern).

[11] Fama-French and momentum factors are obtained from Kenneth French's web site: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[12] Consistent with Cremers and Weinbaum (2010), they report 45 basis points per week in alpha (or 23% annually).

Therefore, informed investors time the horizon by forming expectations of when they believe returns will be realized, which is also observed in portfolio results.

6 Alternative Explanations

Since $PCPD$ can be sensitive to nonsynchronicity of quotes between the puts and calls, or between the option and stock market, microstructure effects may confound the current information-based interpretations. For example, Battalio and Schultz (2006) argues that nonsynchronicity of quotes can be mistaken for deviations from put-call parity. A second confounding factor is that IV-based measures may embed ex-ante volatility information in addition to return information [Ni, Pan and Poteshman (2008), Puhan (2014)]. I specifically address and reject these two concerns.

6.1 Market Microstructure

I rule out alternative microstructure explanations one by one in this subsection. First, I rule out the confounding effects of bid-ask spread in either market ($SBAS$ and $OLBAS$). In columns 1-2 of Table 6.1, I explicitly control for bid-ask spread of the options on the last day of the month as well as the bid-ask spread of the underlying stock, and find that my results are robust to bid-ask spreads. Second, since $PCPD$ takes the difference of implied volatilities of the last call and put traded on the last trading day of each month, the put-call pair should ideally be traded around the same time, if not the same instant. To the extent that liquidity of many equity options can be thin over a day, measurement errors exist in the $PCPD$ measure. To alleviate this concern, I re-run my baseline analysis from Table 4, but including only options that are traded on the last day of the month to avoid stale quotes in options.[13] Doing so reduces the sample size from $273,487$ to $231,972$ observations. Columns 3-4 of Table 10 show that the baseline conclusion is not affected by this filter. The lower diagonal pattern in the 2-by-2 table of coefficients is still

[13] Van Binsbergen, Brandt and Koijen (2012) also measure deviations from put-call parity in *index* options. They match calls and puts that are traded on the same minute of the last trading day of each month. In contrast, some equity options are very limited in volume, but ensuring activity on the last trading day of the month at least tightens the analysis.

clearly observed. Third, quotes may also be nonsynchronous across the option and stock market. Nonsynchronous quotes are well known to induce cross-autocorrelation in assets and potentially across markets [Lo and MacKinlay (1990), Ahn et. al. (2002)], so the predictive power of $PCPD$ can confounded by this. Battalio and Schultz (2006) specifically pointed out that nonsynchronicity in option and stock quotes can lead to erroneous conclusions. They also claim that the option market ceases trading at 4:02pm—2 minutes after the stock market. However, this does not match the Product Specifications of Equity Options on CBOE's website, [14] which states that trading hours of equity options are 8:30am-3pm (CST) [9:30am-4pm (EST)]. Regardless of the ambiguity, my results are robust to an extra-day lag in stock returns[15]. I re-run my analysis based on $PCPD$ predictors constructed on the *second-last* day of the month. This effectively lags all stock returns by a day. Columns 5-6 show that the lower-diagonal pattern survives this 1-day lag. As a result, the empirical observations in this paper cannot be explained by microstructure effects.

6.2 Volatility Effects

Previous studies document that equity options embed volatility information, in addition to returns information. Ni, Pan and Poteshman (2008) find that option market demand for volatility predicts future realized volatility. Puhan (2014) uses straddle positions to investigate volatility information in index options. On the flip side, stocks with different volatility expectations may also have different expected returns in the cross-section. Ang et al. (2006, 2009) document that stocks with higher idiosyncratic volatility have lower cross-sectional returns. Conrad, Dittmar and Ghysels (2013) document the cross-sectional relationship between higher moments of stock returns and expected stock returns. Johnson (2012) uses the VIX term structure to predict S&P 500 returns. I address these volatility-related concerns with two approaches, both of which control for the the *level* of implied volatility. First, I separate forward-looking volatility information from return

[14] https://www.cboe.com/Products/EquityOptionSpecs.aspx

[15] Note that this does not contradict the fact that the predictive power should be strongest in the first few days. Indeed, some previous studies [Chan, Chung and Fong (2002), Chakravarty, Gulen and Mayhew (2004), Pan and Poteshman (2006)] have documented using intraday data that the information gap between the two markets tends to close very quickly even in the first day of trading.

information by controlling for a straddle position consisting of half a call and half a put. $STRAD$ is defined as $(CIV+PIV)/2$. Since the straddle position consists of longing half a call and longing half a put, its payoff will be positive only if stock prices experience big movements in either direction, i.e. a large volatility. Since $RVOL$ is already in my standard controls, adding $STRAD$ also effectively controls for $RVOL - STRAD$, which is investigated by Goyal and Saretto (2009), although their paper has a different focus on option portfolio returns rather than stocks. In Table 11, columns 1-2 include $STRAD$ with standard controls, and horizon timing effects are robust to the straddle. Second, the volatility level can also distort the $PCPD$ measure if put-call parity deviations are generally larger for more volatile options. Columns 3-4 use the *proportional* $PCPD$ ($\%PCPD = PCPD/STRAD$) as regressors, and columns 5-6 include both $\%PCPD$ and $STRAD$ as regressors. Finally, columns 7-8 control for both the straddle and bid-ask spreads from Section 6.1. In summary, volatility effects do not confound the horizon timing effects.

7 Conclusion

My contribution in this paper is to document robust evidence that horizon timing effects exist in equity option market. I find that long-dated options embed both short and long-horizon information despite the availability of a cheaper short-dated option, while short-dated options mainly embed short-horizon information, but can also contain information beyond their expiration dates. This forms a lower-diagonal pattern discussed in the paper. These results are especially consistent with the existence of uncertainty in price discovery speed or timing, and that informed investors not only take this uncertainty into consideration, but are also successful in their timing prediction. My results are also consistent with information-based interpretations because the strength of the predictability trends in an expected fashion in accordance with existing asymmetric information models for information flow between the stock and option markets.

The current motivating model in Section 2 and Appendix B is quite simple, and some theoretical extensions are possible. First, risk-neutrality assumption can be relaxed. One challenge is that common normality assumption (e.g. with a CARA utility) is not reasonable due to the non-linear payoff structure of options. Some previous works have tackled this challenge with various levels of complexities [Liu and Pan (2003), Jones (2006), Faias and Santa-Clara (2011), Eraker (2013)]. Second, it may be interesting to combine features of the horizon timing effects with existing information asymmetry models. In particular, my simplified setting currently takes transaction cost and option price as exogenously given, but investment activities of the traders can of course endogenously affect both price and liquidity [Kyle (1985), Back (1992)]. Thus, closing the model with equilibrium considerations may be a promising avenue of future work.

Appendix A List of Acronyms

Acronym	Meaning or Definition
ATM/ITM/OTM	At-the-money/In-the-money/Out-of-the-money
CIV	Call Implied Volatility
PIV	Put Implied Volatility
R_N	Nth-month Stock Return
R_{N-M}	Nth to Mth-month Stock Return
$PCPD$	Put-call Parity Deviation = $PIV_{ATM} - CIV_{ATM}$
$PCPD_{T_2 \perp T_1}$	$PCPD_{T_2}$ orthogonalized to $PCPD_{T_1}$
$\Delta PCPD$	Change in $PCPD$ Over 1 Month
$\%PCPD$	$PCPD/STRAD$
$STRAD$	$(CIV + PIV)/2$
$SMIRK$	$PIV_{OTM} - CIV_{ATM}$
$PSKEW$	$PIV_{OTM} - PIV_{ATM}$
O/S	Monthly Option-to-stock Volume Ratio
$O/SHROUT$	Monthly Option Volume / Num. of Shares Outstanding
$BETA$	Monthly Betas
$SIZE$	Log of Market Capitalization
$B2M$	Book-to-Market Ratio
MOM	Past 2-12 Month Momentum
Short REV	Past Month Stock Return
$ILLIQ$	Amihud's illiquidity (Monthly Stock Return/Monthly Volume)
$RVOL$	Realized Volatility (Sum of squared daily stock returns)
V_c/V_p	Call/Put Volume Ratio
OI_c/OI_p	Call/Put Open Interest Ratio
$SBAS$	Stock Bid-ask Spread
$\%SBAS$	BAS / Mid Price
$OBAS$	Option Bid-ask Spread
$OLBAS$	Option Bid-ask Spread on last trading day of the month
$\%OBAS$	$OBAS$ / Mid Premium

Appendix B A Simple Model of Horizon Timing

Using the notations and simplifying assumptions described in Section 2, I will derive the optimal investment decision using a backward induction approach. I enforce the following participation constraint to ensure that it is profitable for the informed agent to act on her information, at least when payoff is certain in the short-horizon:

$$\text{Participation Constraint:} \quad \Delta_1 R \geq c_1 + F_1$$

Since the informed trader is risk-neutral, I only maximize the expected payoff of each scenario. Risk-neutrality also implies that a mixed portfolio of two or more investment alternatives is never optimal, except on the indifference boundary. At $t = 0$, by backward induction, the first-order terms in the risk-neutral agent's expected payoff (π_0) is:

$$\pi_0 = \begin{cases} \frac{W}{c_2+F_2}[p\Delta_2 R + p(c_1 - F_1) + (1-p)\Delta_1 R] & \text{if invest in long-dated option at } t = 0 \\ \frac{W}{c_1+F_1} p\Delta_1 R + 0 & \text{if invest in short-horizon option at } t = 0 \\ 0 \qquad\qquad + pW + (1-p)\frac{W}{c_1+F_1}\Delta_1 R & \text{if decide to wait till } t = T_1 \end{cases}$$

For simplicity, I normalize to unit wealth $W = 1$ for this analysis. In the first case, the agent invests in $\frac{1}{c_2+F_2}$ long-dated option at $t = 0$, where F_2 is the one-way transaction cost, for each option traded, in return for an expected payoff of $p\Delta_2 R$ in the first period. At $t = T_1$, if the payoff did not occur in the first period, she is then certain that the payoff will come in the second period, and so will continue to hold the same option. However, if the payoff occurred in the first period, she would want to sell the option (as a short-dated option at a price of $c_1 - F_1$) to reap the remaining time-value of the option.[16] The second case is similar except that if the payoff did not occur in the first period (with probability $1 - p$) and the option expires worthless, the agent would have exhausted her wealth at $t = T_1$. The third case is the simplest in that the agent will only

[16] A *risk-averse* agent would also want to take profit and sell the option at $t = T_1$ to avoid extra risk, conditional on the payoff having already been realized in the first period.

invest in the short-horizon option in the second period, conditional on the payoff not arriving in the first period. The participation constraint ensures that it is profitable to do so. If the payoff occurs in the first period, then the agent would simply hold the wealth W without investing in anything for both periods.

Since my empirical strategy currently involves ATM options only, I restrict attention to ATM options in the current analysis, which I define as $\Delta_1 = \Delta_2 = 0.5$. Define $\gamma \equiv \frac{c_1+F_1}{c_2+F_2}$ to be the ratio of short-dated to long-dated option premiums, including the one-way transaction cost incurred. Since premium and transaction cost of long-dated options are higher ($c_2 > c_1$ and $F_2 > F_1$), therefore $\gamma < 1$. The optimal investment decision can be solved as a function of these parameters: p, R, c_1, γ, F_1, and F_2. I solve for the parameter boundaries below:

- The agent prefers long-dated option to waiting if $p > \frac{\frac{1}{2}R(1-\gamma)}{\frac{1}{2}R-\gamma(c_2+F_2-c_1+F_1)} \equiv P_{low}$.
- The agent prefers short-dated to long-dated option if $p > \frac{\frac{1}{2}R\gamma}{\frac{1}{2}R-\gamma(c_1-F_1)} \equiv P_{high}$.
- The agent prefers short-dated to waiting if $p > \frac{\frac{1}{2}R}{R-(c_1+F_1)} \equiv P_{med}$.

The following results are easily derived with some algebra:

Result 1: When payoff R is sufficiently large (specifically $R > R^* \equiv \frac{2\gamma}{1-\gamma}(c_1 - F_1)$) and $\gamma > \frac{1}{2}$, three regions of investment decisions are created. This result can be proven by imposing $0 < P_{low} < P_{med} < P_{high} < 1$, and then simplify each inequality. This result motivates the stylized facts and empirical observations in this paper. (The core intuition of this result is described in Section 2 and is not repeated here.)

Result 2: Ignoring transaction costs of options ($F_1 = 0, F_2 = 0$), $P_{high} \to \frac{c_1}{c_2}$ as $R \to \infty$ and $P_{low} \to 1 - \frac{c_1}{c_2}$ as $R \to \infty$. This result can be proven easily from the expressions of P_{high} and P_{low}, and their intuitive interpretations are discussed in Section 2.

Appendix C Persistence of Three Documented Predictors

The persistence of three documented IV-based predictors of stock returns are compared. Cremers and Weinbaum (2010) document that higher $PCPD$ predicts a higher stock return. Xing, Zhang and Zhao (2010) find that higher $SMIRK$ predicts a lower stock return. An et. al. (2014) document that higher $\Delta PCPD$ predicts higher stock return. The definitions of these predictors are given in Appendix A. Using Fama-Macbeth regressions as described in Section 4, I test the persistence of each 6-month predictor, by regressing R_N ($N = 1, 2, \cdots, 6$) on each predictor alone with an intercept. The regressors are normalized to have zero-mean and unit standard deviation to facilitate comparisons. For this analysis, 6-month options are chosen to give the predictor the best chance of having a persistent predictability, but the conclusions about their relative persistence remain unchanged even with options of other maturities, and with or without controls. The table below display the results. First, the statistical significance of all three documented predictors are confirmed for the R_1 horizon. Second, $PCPD_6$ is the most persistent, with statistically significance throughout the next 6 months. $SMIRK_6$ also seems persistent, but its persistence actually comes from $PCPD_6$, since $PSKEW_6$ is not persistent as shown in the third row. $\Delta PCPD$ is also not persistent as well. This actually does not contradict An et. al. (2014), since they use R_{1-N} as dependent variable as opposed to R_N. In other words, most of the predictability of $\Delta PCPD$ on R_{1-N} actually comes from R_1. As a result, $PCPD$ is the only measure that is persistent enough for a decomposition of horizon timing effects up to a 6-month period.

Dependent Var.	R_1	R_2	R_3	R_4	R_5	R_6
	(1)	(2)	(3)	(4)	(5)	(6)
$PCPD_{6,ATM}$	0.46***	0.17***	0.20***	0.12**	0.12**	0.10**
	(7.1959)	(3.9497)	(4.9368)	(2.5758)	(2.395)	(2.1298)
$SMIRK_6$	-0.43***	-0.17***	-0.16**	-0.14**	-0.10	-0.0739
	(-4.8168)	(-2.5702)	(-2.2859)	(-2.0817)	(-1.4858)	(-1.0718)
$PSKEW_6$	-0.0110	-0.0534	0.0031	-0.0351	0.0031	-0.0078
	(-0.15182)	(-0.67061)	(0.040622)	(-0.51026)	(0.047101)	(-0.11384)
$\Delta PCPD_{6,ATM}$	0.26***	-0.0215	0.0622	0.0184	0.0061	0.0154
	(5.8332)	(-0.58509)	(1.4438)	(0.58475)	(0.15557)	(0.41835)
Controls	No	No	No	No	No	No
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487

References

Abel, A. B., J. C. Eberly, and S. Panageas, 2013, Optimal Inattention to the Stock Market with Information Costs and Transaction Costs, Econometrica 81(4), 1455-1481.

Abel, A. B., J. C. Eberly, and S. Panageas, 2007, Optimal Inattention to the Stock Market, American Economic Review 97(2).

Ahn, D. H., J. Boudoukh, M. Richardson, and R. F. Whitelaw, 2002, Partial Adjustment or Stale Prices? Implications from Stock Index and Futures Return Autocorrelations, Review of Financial Studies 15(2), 655-689.

Alvarez, F., L. Guiso and F. Lippi, 2013, Durable Consumption and Asset Management with Transaction and Observation Costs, American Economic Review.

An, B. J., A. Ang, T. G. Bali, and N. Cakici, 2014, The Joint Cross Section of Stocks and Options, Journal of Finance (forthcoming).

Andersen, T., T. Bollerslev, F. X. Diebold, and C. Vega, 2003, Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange, American Economic Review 93(1), 38-62.

Andersen, T., T. Bollerslev, F. X. Diebold, and C. Vega, 2007, Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets, Journal of International Economics 73, 251-277.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, The cross-section of volatility and expected returns, Journal of Finance 61, 259-299.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2009, High idiosyncratic volatility and low returns: International and further U.S. evidence, Journal of Financial Economics 91, 1-23.

Back, K., 1992, Insider Trading in Continuous Time, Review of Financial Studies 5, 387-409.

Badrinath, S. G., J. R. Kale, and T. H. Noe, 1995, Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns, Review of Financial Studies 8, 401-430.

Battalio, R. and P. Schultz, 2006, Options and the bubble, Journal of Finance 61, 2071-2102.

Bernard, V. L., J. K. Thomas, 1989, Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?, Journal of Accounting Research, 27, 1-36.

Bernard, V. L., J. K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, Journal of Accounting and Economics 13(4), 305-340.

van Binsbergen, Jules H., Michael W. Brandt, and Ralph S.J. Koijen, 2012, On the Timing and Pricing of Dividends, American Economic Review, 102(4): 1596-1618. Winner of the 2010 Outstanding Paper Award of the Swiss Finance Institute

Brennan, M. J., N. Jagadeesh, and B. Swaminathan, 1993, Investment Analysis and the Adjustment of Stock Prices to Common Information, Review of Financial Studies 6, 799-824.

Cai, J., R. Y. K. Ho, R. Korajczyk, and Z. Zhang, 2013, Price Discovery, Foreign Ownership, and Rule of Law, Working paper, Northwestern University.

Cao, C., Z. Chen, and J. M. Griffin, 2005, The informational content of option volume prior to takeovers, Journal of Business 78, 1073-1109.

Chakravarty, S., H. Gulen, and S. Mahew, 2004, Informed trading in stock and option markets, Journal of Finance 59, 1235-1257.

Chen, Z. and A. Lu, 2013, Slow Diffusion of Information and Price Momentum in Stocks: Evidence from Options Markets, Working Paper.

Chan, K., Y. P. Chung, and W.-M. Fong, 2002, The Informational Role of Stock and Option Volume, Review of Financial Studies, 15, 1049-1075.

Chordia, T., and B. Swaminathan, 2000, Trading Volume and Cross-Autocorrelations in Stock Returns, Journal of Finance 55, 913-935.

Cox, J., S. Ross, and M. Rubinstein, 1979, Option Pricing: A Simplified Approach, Journal of Financial Economics 7, 229-263.

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.

Duffie, D., 2010, Asset Price Dynamics with Slow-Moving Capital, AFA Presidential Address, ASSA Meetings Atlanta January 2-6 2010.

Easley, D., M. O'Hara, and P. Srinivas, 1998, Option Volume and Stock Prices: Evidence on Where Informed Traders Trade, Journal of Finance, 53, 431-465.

Eraker, B., 2013, The performance of model based option trading strategies, Review of Derivative Research 16, 1-23.

Faias, J. A., and P. Santa-Clara, 2011, Optimal Option Portfolio Strategies, Working Paper.

Goyal, A., and A. Saretto, 2009, Cross-section of option returns and volatility, Journal of Financial Economics 94, 310-326.

Hou, K., 2007, Industry Information Diffusion and the Lead-lag Effect in Stock Returns, Review of Financial Studies 20(4), 1113-1138.

Jegadeesh, N., and J. Livnat, 2006, Post-Earnings-Announcement Drift: The Role of Revenue Surprises, Financial Analysts Journal 62(2), 22-34.

Jones, C., 2006, A nonlinear factor analysis of S&P 500 index option returns, Journal of Finance 61, 2325-2363.

Kyle, A., 1985, Continuous auctions and insider trading. Econometrica 53, pp. 1315-1336.

Lee, J., C.H. Yi, 2001, Trade Size and Information-Motivated Trading in the Options and Stock Markets, Journal of Financial and Quantitative Analysis, 36(4), 485-501.

Lin, T.C., X. Liu, and J. Driessen, 2013, Why Do Option Prices Predict Stock Returns?, Working Paper.

Liu, J., and J. Pan, 2003, Dynamic derivative strategies, Journal of Financial Economics 69, 401-430.

Lo, A., and MacKinlay, A. C., 1990, An Econometric Analysis of Nonsynchronous Trading, Journal of Econometrics 45, 181-211.

Jegadeesh, N. and S. Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. Journal of Finance, 48: 65-91.

Johnson, T. L., 2012, Equity Risk Premia and the VIX Term Structure, Working Paper.

Johnson, T. L., and E. C. So, 2012, The option to stock volume ratio and future returns, Journal of Financial Economics 106 (2): 262-286.

Kehrle, Kerstin and Puhan, Tatjana-Xenia, 2013, The Information Content of Option Demand, Swiss Finance Institute Research Paper No. 12-43.

Newey, W. and K. West, 1987, A simple positive, semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 29, pp. 229-256.

Ni, S. X., Jun Pan, and A. M. Poteshman, 2008, Volatility information trading in the option market, Journal of Finance 63 (3): 1059-1091.

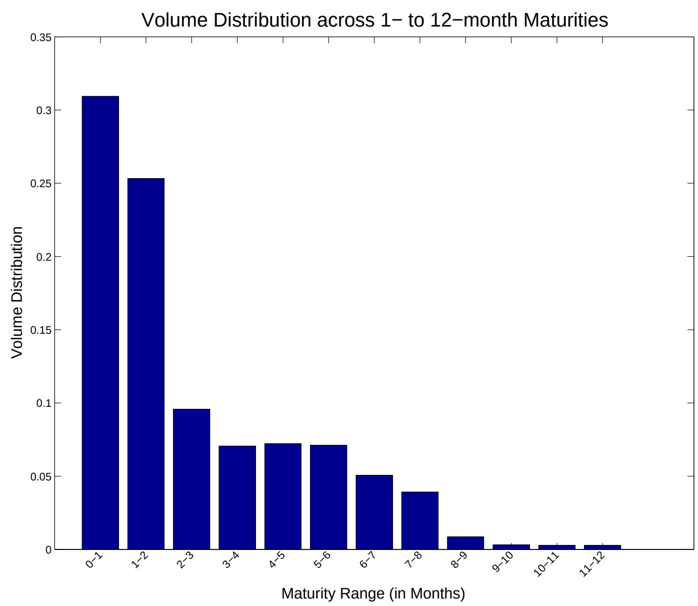
Pan, Jun, and Allen M. Poteshman, 2006, The information in option volume for future stock prices, Review of Financial Studies 19, 871-908.

Roll, R., E. Schwartz, and A. Subrahmanyam, 2009, Options trading activity and firm valuation, Journal of Financial Economics 94, 345-360.

Stoll, H. R., 1969, The relationship between put and call option prices, Journal of Finance 24(5): 801-824.

Xing, Y., X. Zhang, R. Zhao, 2010, What Does Individual Option Volatility Smirk Tell Us About Future Equity Returns? Journal of Financial and Quantitative Analysis.

Figure 1: **Volume Distribution and Bid-ask Spreads across Different Option Maturities.** The top figure plots the option volume distribution across 1-12 month maturities. The bottom figure plots the ratio of median bid-ask spreads between Nth month and the first month. Both graphs are first computed for each firm-date, then averaged over all firm-dates. Both graphs indicate that the liquidity decreases in maturity. Options beyond eight months have very little liquidity, making them prohibitive to trade.



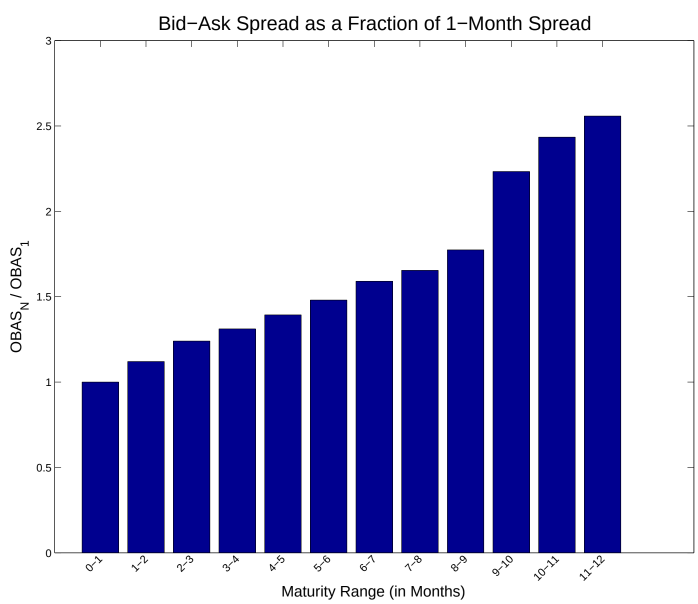


Figure 2: $PCPD$ **and** $SMIRK$ **Measures on a Typical Volatility Smile**

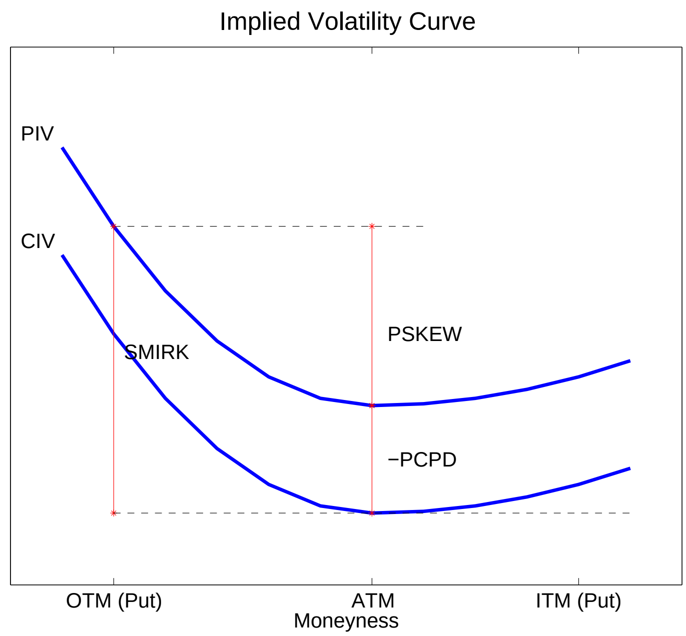


Table 1: **Number of firms and firm-month observations by year in the merged dataset.**

Year	Number of Firms	Number of Observations
1996	1,270	11,474
1997	1,561	15,036
1998	1,714	16,544
1999	1,690	16,413
2000	1,610	14,429
2001	1,515	14,105
2002	1,586	15,992
2003	1,538	15,913
2004	1,646	17,119
2005	1,707	18,000
2006	1,813	18,744
2007	1,839	19,967
2008	1,925	19,727
2009	1,854	19,026
2010	1,920	20,248
2011	2,028	20,750
Average	1,701	17,093
Total		273,487

Table 2: **Summary Statistics of Option-based Measures across Different Maturities.**
Panel A is based on the implied volatility of at-the-money calls (CIV) and puts (PIV) with 1, 3, and 6 months to expiration. Panel B is based on the put-call parity deviation ($PCPD = CIV - PIV$) measure computed using options of 3 and 6 months to expiration. $PCPD_{6\perp 3}$ is $PCPD_6$ orthogonalized with respect to $PCPD_3$, computed as the residuals of regressing $PCPD_6$ on $PCPD_3$ with an intercept. Therefore, the zero correlation between $PCPD_3$ and $PCPD_{6\perp 3}$ is by construction. All summary statistics are computed based on the entire panel, first in the cross-section and then averaged over time. Mean, standard deviation, skewness, median, and mean absolute deviation (MAD) are reported in the bottom rows. MAD is omitted from Panel A, because CIV and PIV are positive values by definition.

Panel A: Summary Statistics of CIV and PIV across Different Maturities

	CIV_1	CIV_3	CIV_6	PIV_1	PIV_3	PIV_6
CIV_1	1.00	0.94	0.93	0.92	0.91	0.89
CIV_3	0.94	1.00	0.98	0.91	0.95	0.94
CIV_6	0.93	0.98	1.00	0.90	0.94	0.96
PIV_1	0.92	0.91	0.90	1.00	0.95	0.93
PIV_3	0.91	0.95	0.94	0.95	1.00	0.98
PIV_6	0.89	0.94	0.96	0.93	0.98	1.00
Mean	0.47	0.46	0.45	0.48	0.47	0.46
Std	0.19	0.18	0.17	0.19	0.18	0.17
Skewness	1.56	1.17	1.03	1.72	1.33	1.18
Median	0.44	0.43	0.43	0.45	0.44	0.43

Panel B: Summary Statistics of $PCPD$ across Different Maturities

	$PCPD_3$	$PCPD_6$	$PCPD_{6\perp 3}$
$PCPD_3$	1.00	0.81	0.00
$PCPD_6$	0.81	1.00	0.56
$PCPD_{6\perp 3}$	0.00	0.56	1.00
Mean	-0.0072	-0.0075	0.00
Std	0.0533	0.0471	0.0264
Skewness	-1.80	-2.61	0.40
Median	-0.0048	-0.0047	0.000681
MAD	0.0260	0.0227	0.0120

Panel C: Summary Statistics of Bid-ask Spreads in Stocks and Options

	$SBAS$	$OBAS$	$\%SBAS$	$\%OBAS$
$SBAS$	1.00	0.20	0.35	-0.14
$OBAS$	0.20	1.00	-0.0672	0.0915
$\%SBAS$	0.35	-0.0672	1.00	0.39
$\%OBAS$	-0.14	0.0915	0.39	1.00
Mean	0.16	0.67	0.0064	0.51
Std	0.12	4.61	0.0047	0.18
Skewness	4.18	7.58	3.39	0.59
Median	0.13	0.37	0.0051	0.49

Table 3: **Documented Predictability of $PCPD$.**
Coefficients and t-statistics are reported based on Fama-Macbeth regressions of 1-month forward return on the respective regressor(s) with an intercept. All t-statistics are adjusted with Newey-West adjustment of 6 lags. The standard control variables used in this paper include market beta ($BETA$), log of market capitalization ($SIZE$), log of book-to-market ratio ($B2M$), past 2-to-12 month momentum (MOM), Amihud's illiquidity ($ILLIQ$), 1-month past return to proxy for short-term reversal (Short REV), realized volatility ($RVOL$) computed as the sum of squared daily returns over the past month, left-skew computed based on puts ($PSKEW = PIV_{OTM} - PIV_{ATM}$), call/put volume ratio (V_c/V_p), and call/put open-interest ratio (OI_c/OI_p). R^2 is reported at the bottom. 10%/5%/1% significance levels are marked with */**/***.

Dependent Var.	R_1 (1)	R_1 (2)	R_{1-3} (3)	R_{1-3} (4)
$PCPD_{1,ATM}$	5.90***	5.06***		
	(7.0125)	(6.1089)		
$PCPD_{3,ATM}$			4.69***	4.31***
			(7.7013)	(7.5645)
BETA		-0.0339		0.0932
		(-0.20452)		(0.95734)
SIZE		-0.0749		-0.0480
		(-1.0254)		(-0.78162)
B2M		0.12		0.13
		(0.99007)		(1.1374)
MOM		0.0028		0.0019
		(0.75072)		(0.59559)
ILLIQ		0.13		0.0425
		(0.77114)		(0.374)
Short REV		-0.0189**		-0.0061
		(-2.4766)		(-1.05)
RVOL		-0.14*		-0.14*
		(-1.8651)		(-1.9118)
$PSKEW_3$		-1.10		-1.49
		(-0.88783)		(-1.4969)
V_c/V_p		0.0021		-0.000970
		(0.92353)		(-0.81235)
OI_c/OI_p		-0.0011		-0.0033
		(-0.26508)		(-1.074)
Controls	No	Yes	No	Yes
Intercept	Yes	Yes	Yes	Yes
Normalized Coeff.	No	No	No	No
R^2	0.30%	9.11%	0.32%	9.02%
# Obs.	273487	273487	273487	273487

Table 4: **Baseline Results—Horizon Timing Effect Based on 3/6-month $PCPD$.**
In the first two columns, I run Fama-Macbeth regressions of 1-3 and 4-6 month forward return respectively on the two regressors with an intercept. $PCPD_{6\perp3}$ means that $PCPD_{6,ATM}$ is orthogonalized with respect to $PCPD_{3,ATM}$, so $PCPD_{3,ATM}$ is interpreted as the base. In the third and fourth column, I add standard controls to the regression. In the fifth and sixth column, same regressions are done except the regressors are standardized to have zero mean and unit standard deviation for better comparability across coefficients. Panel B is the result of Fama-Macbeth regressions of the L^{th} month forward return ($L = 1, 2, 3, 4, 5, 6$) on the regressors with an intercept without controls. Similar pattern also exists with the standard controls (not reported). R^2 is reported on at the bottom. 10%/5%/1% significance levels are marked with */**/***.

Panel A: Horizon Timing Effect Grouped by 3-month Returns

Dependent Var.	R_{1-3} (1)	R_{4-6} (2)	R_{1-3} (3)	R_{4-6} (4)	R_{1-3} (5)	R_{4-6} (6)
$PCPD_{3,ATM}$	4.69***	0.99	4.29***	1.32***	0.22***	0.0602**
	(7.7013)	(1.5563)	(7.5063)	(2.6151)	(7.8549)	(2.384)
$PCPD_{6\perp3}$	6.14***	6.16***	5.06***	5.26***	0.13***	0.12***
	(4.3982)	(4.5831)	(4.6489)	(4.6745)	(5.0378)	(5.0668)
Controls	No	No	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	No	No	No	No	Yes	Yes
R^2	0.50%	0.39%	9.16%	8.50%	9.16%	8.50%
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487

Panel B: Horizon Timing Effect Month-by-month

Dependent Var.	R_1 (1)	R_2 (2)	R_3 (3)	R_4 (4)	R_5 (5)	R_6 (6)
$PCPD_{3,ATM}$	0.42***	0.0959***	0.14***	0.0531	0.0540	0.0734**
	(7.7058)	(3.0013)	(3.5355)	(1.4174)	(1.4126)	(2.159)
$PCPD_{6\perp3}$	0.13***	0.13***	0.12***	0.11***	0.13***	0.12***
	(3.5393)	(3.9326)	(4.0487)	(3.584)	(3.4804)	(3.7769)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes
R^2	9.30%	8.53%	8.13%	8.05%	7.62%	7.50%
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487

Table 5: **Horizon Timing Effect Based on 2/4/6-month $PCPD$.**
In the first two columns, I run Fama-Macbeth regressions of 1-2, 3-4, and 5-6 month forward return respectively on the three regressors with an intercept. $PCPD_{6\perp 4,2}$ is computed by taking the residual of regressing $PCPD_6$ on $PCPD_4$ and $PCPD_2$ with an intercept. In columns 4-6, I add standard controls to the regression. In columns 7-9, same regressions are done except the regressors are standardized to have zero mean and unit standard deviation for better comparability across coefficients. Panel B is the result of Fama-Macbeth regressions of the L^{th} month forward return ($L = 1, 2, 3, 4, 5, 6$) on the regressors with an intercept without controls. Similar pattern also exists with the standard controls (not reported). R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

Panel A: Horizon Timing Effect Grouped by 2-month Returns

Dependent Var.	R_{1-2} (1)	R_{3-4} (2)	R_{5-6} (3)	R_{1-2} (4)	R_{3-4} (5)	R_{5-6} (6)	R_{1-2} (7)	R_{3-4} (8)	R_{5-6} (9)
$PCPD_{2,ATM}$	4.25***	0.90	1.23***	3.79***	0.75	1.57***	0.23***	0.0551*	0.0670**
	(8.452)	(1.2159)	(2.6416)	(7.4446)	(1.3002)	(3.4491)	(7.4286)	(1.7414)	(2.501)
$PCPD_{4\perp 2}$	4.01***	4.48***	0.88	3.79***	3.88***	1.56	0.13***	0.12***	0.0476*
	(2.9158)	(3.3168)	(0.6461)	(3.0332)	(3.3913)	(1.322)	(4.1135)	(4.4553)	(1.7747)
$PCPD_{6\perp 4,2}$	6.56***	9.19***	8.12***	5.48***	7.60***	6.70***	0.10***	0.10***	0.11***
	(3.2005)	(4.069)	(3.6318)	(3.0793)	(4.0024)	(3.6469)	(3.9933)	(4.3033)	(4.1621)
Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	No	No	No	No	No	No	Yes	Yes	Yes
R^2	0.59%	0.53%	0.51%	9.36%	8.64%	8.13%	9.36%	8.64%	8.13%
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487	273,487	273,487	273,487

Panel B: Horizon Timing Effect Month-by-month

Dependent Var.	R_1 (1)	R_2 (2)	R_3 (3)	R_4 (4)	R_5 (5)	R_6 (6)
$PCPD_{2,ATM}$	0.38***	0.0693**	0.0937**	0.0165	0.0543	0.0798**
	(6.9984)	(2.0057)	(2.4405)	(0.34923)	(1.5582)	(2.1586)
$PCPD_{4\perp 2}$	0.18***	0.0887***	0.14***	0.0914***	0.0529	0.0422
	(4.0113)	(2.7057)	(3.8299)	(2.7818)	(1.5205)	(1.3146)
$PCPD_{6\perp 4,2}$	0.0838**	0.12***	0.0982***	0.11***	0.11***	0.11***
	(2.5832)	(3.9164)	(3.5671)	(3.3096)	(2.7257)	(3.4703)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes
R^2	9.41%	8.62%	8.22%	8.18%	7.75%	7.59%
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487

Figure 3: **Decay of predictive power of $PCPD$ over 12-month horizon.**
Each point on each line correspond to a Fama-Macbeth regression of T-month cumulative forward return on each regressor standalone (e.g. $PCPD_{1,ATM}$). The eight lines correspond to using predictors computed based on options with eight different maturities: 1, 2, 3, 4, 5, 6, 9, and 12 months respectively.

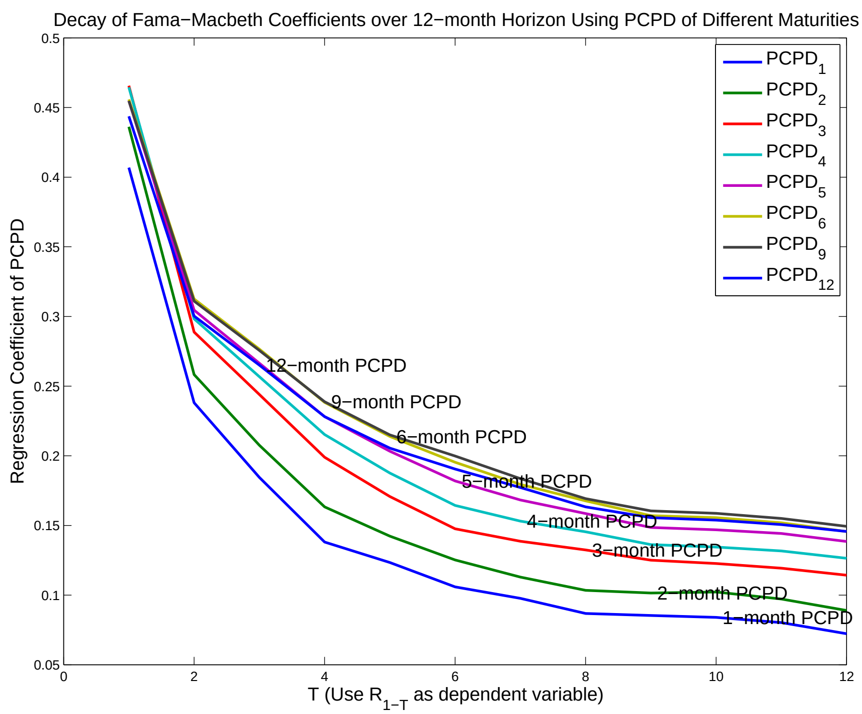


Table 6: **Evidence of Information-based Explanation: Sort on Option-to-stock Volume Ratio.** I sort on option/stock volume ratio (O/S) in each cross-section of Fama-MacBeth regression. This result confirms that O/S is an effective conditioning variable. It strengthens the information-origin of horizon timing effect. R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

Dependent Var.	R_{1-3} (1)	R_{1-3} (2)	R_{1-3} (3)	(4)	R_{4-6} (5)	R_{4-6} (6)	R_{4-6} (7)	(8)
$PCPD_{3,ATM}$	0.21***	0.25***	0.28***	0.0738	-0.0236	0.0638	0.15**	0.17***
	(4.6603)	(5.6086)	(4.7324)	(0.93075)	(-0.72187)	(1.4411)	(2.4444)	(2.651)
$PCPD_{6\perp 3}$	0.11***	0.15***	0.16***	0.0463	0.0766*	0.15***	0.16***	0.0789*
	(3.2333)	(3.7413)	(3.0567)	(0.78835)	(1.9491)	(3.6701)	(4.1404)	(1.842)
Controls	No	No	No		No	No	No	
Intercept	Yes	Yes	Yes		Yes	Yes	Yes	
Normalized Coeff.	Yes	Yes	Yes		Yes	Yes	Yes	
R^2	0.92%	0.88%	1.33%		0.77%	0.97%	1.06%	

Table 7: **Evidence of Information-based Explanation: Sort on Turnover Measures.**
In Panel A, I sort on option turnover ($O/SHROUT$), defined as the ratio of option volume to number of shares outstanding, in each cross-section of the Fama-MacBeth regression. In Panel B, I sort on stock turnover ($S/SHROUT$), which I argue is ambiguous in Section 4.3.2. This result confirms that $O/SHROUT$ is an effective conditioning variable, but $S/SHROUT$ is not. In particular, Panel A strengthens the information-origin of horizon timing effect. R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

Panel A: Sorting on $O/SHROUT$

Dependent Var.	R_{1-3} (1)	R_{1-3} (2)	R_{1-3} (3)	(4)	R_{4-6} (5)	R_{4-6} (6)	R_{4-6} (7)	(8)
$PCPD_{3,ATM}$	0.21***	0.23***	0.29***	0.0852	-0.0006	0.0608	0.11*	0.11*
	(5.2462)	(5.8161)	(5.5244)	(1.3065)	(-0.021001)	(1.3592)	(1.8396)	(1.6968)
$PCPD_{6\perp 3}$	0.10***	0.13***	0.16***	0.0554	0.0500*	0.15***	0.16***	0.11**
	(2.9374)	(4.0164)	(3.6697)	(1.121)	(1.7126)	(3.3278)	(3.5026)	(2.4011)
Controls	No	No	No		No	No	No	
Intercept	Yes	Yes	Yes		Yes	Yes	Yes	
Normalized Coeff.	Yes	Yes	Yes		Yes	Yes	Yes	
R^2	0.94%	0.93%	1.09%		0.81%	1.04%	0.96%	

Panel B: Sorting on $S/SHROUT$ (Should be Ambiguous)

Dependent Var.	R_{1-3} (1)	R_{1-3} (2)	R_{1-3} (3)	(4)	R_{4-6} (5)	R_{4-6} (6)	R_{4-6} (7)	(8)
$PCPD_{3,ATM}$	0.22***	0.25***	0.24***	0.0230	0.0671*	0.0574	0.0332	-0.0338
	(6.8053)	(6.5283)	(5.1425)	(0.47457)	(1.7763)	(1.551)	(0.56391)	(-0.54315)
$PCPD_{6\perp 3}$	0.11***	0.16***	0.16***	0.0531	0.0654**	0.10***	0.20***	0.13**
	(3.3446)	(3.8021)	(4.3694)	(1.1496)	(2.1751)	(3.1178)	(3.4603)	(2.1408)
Controls	No	No	No		No	No	No	
Intercept	Yes	Yes	Yes		Yes	Yes	Yes	
Normalized Coeff.	Yes	Yes	Yes		Yes	Yes	Yes	
R^2	1.18%	0.96%	0.87%		1.10%	0.78%	0.83%	

Table 8: **Robustness Tests of Horizon Timing Effect.**
This table presents various robustness checks. Columns 1-2 trim outliers of PIV and CIV at 1% and 99% levels. Columns 3-4 include all low-priced stocks below $5. Columns 5-6 and 7-8 presents individual sub-period analysis for 1996-2003 and 2004-2011 respectively. Columns 9-10 alters the orthogonalization procedure to a simple subtraction $PCPD_6 - PCPD_3$ of signals. Columns 11-12 include the controls in the orthogonalization procedure. The horizon timing effect and lower-diagonal pattern are robust to all different specifications. R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

	Trim PIV and CIV		Include Low-priced		Subperiod 96-03		Subperiod 04-11		Orth by Diff		Orth to Controls	
Dependent Var.	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$PCPD_{3,ATM}$	0.21***	0.0746***	0.22***	0.0692***	0.28***	0.0412	0.16***	0.0790**	0.30***	0.13***	0.22***	0.0601**
	(7.1369)	(3.3006)	(8.5497)	(2.8108)	(6.6942)	(1.1471)	(5.479)	(2.2517)	(7.9005)	(3.9943)	(7.8769)	(2.366)
$PCPD_{6\perp 3}$	0.12***	0.0909***	0.13***	0.10***	0.19***	0.16***	0.0665***	0.0788***				
	(4.3883)	(3.7089)	(4.1869)	(3.4692)	(4.8421)	(4.2931)	(2.84)	(3.1641)				
$PCPD_6 - PCPD_3$									0.15***	0.14***		
									(5.1795)	(5.0685)		
$PCPD_{6\perp 3,controls}$											0.12***	0.11***
											(5.0676)	(5.1318)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	9.09%	8.37%	9.32%	8.67%	11.15%	10.46%	7.19%	6.56%	9.16%	8.50%	9.16%	8.50%
# Obs.	267072	267072	285675	285675	119906	119906	153581	153581	273,487	273,487	273,487	273,487

Table 9: **Horizon Timing Effects using Portfolio Sorts.**
The column header in the top row indicates the variable on which sorting is done. S/H represents the number of months to skip (S) and number of months to hold the portfolio (H) respectively. For example, $S = 3, H = 3$ means skip 3 months and hold 3 months, so it is analogous to a 4-6 month forward return. I sort firms into 10 deciles, and report portfolio alphas for each decile, controlled for the Fama-French three factors and momentum factor. The bottom row shows the portfolio alpha resulted by longing decile 10 (high predictor value) and shorting decile 1 (low predictor value). All portfolio alphas are annualized and already in percentages. 10%/5%/1% significance levels are marked with */**/***.

Sort on:	$PCPD_1$	$PCPD_3$		$PCPD_{6\perp 3}$	
Skip/Hold	$S = 0, H = 1$	$S = 0, H = 3$	$S = 3, H = 3$	$S = 0, H = 3$	$S = 3, H = 3$
	(1)	(2)	(3)	(4)	(5)
(1)	-9.39	-4.63	1.14	-1.66	0.82
(2)	-3.26	0.36	4.07	1.75	2.97
(3)	-3.69	1.18	3.51	2.08	3.77
(4)	1.80	1.91	3.01	2.43	3.62
(5)	1.70	2.19	3.13	2.67	3.30
(6)	2.77	3.30	4.02	3.05	3.78
(7)	3.60	4.24	3.38	2.44	3.64
(8)	4.50	4.69	4.68	4.11	4.72
(9)	7.90	5.15	4.47	3.80	4.55
(10)	12.39	7.06	4.77	4.76	5.02
Diff	21.78	11.69	3.63	6.42	4.20
	(6.67***)	(6.82***)	(1.93*)	(5.01***)	(2.87***)

Table 10: **Rejecting Microstructure Effects.**
This table presents evidence to rule out microstructure effects. Columns 1-2 control for stock and option bid-ask spreads. Columns 3-4 avoid stale quotes by including only options that have been traded on the last trading day of the month. Columns 5-6 construct $PCPD$ measures from the second-last trading day of the month, effectively providing a 1-day lag in stock returns. All these specifications control for microstructure effects. The horizon timing effects and the lower-diagonal pattern in the coefficients survive all tests. R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

	Control for Spreads		Avoid Stale Quotes		1-day Lag in Stock Returns	
Dependent Var.	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}	R_{1-3}	R_{4-6}
	(1)	(2)	(3)	(4)	(5)	(6)
$PCPD_{3,ATM}$	0.22***	0.0598**	0.24***	0.0564*	0.18***	0.0596**
	(7.9005)	(2.3581)	(7.8106)	(1.8755)	(6.375)	(2.5681)
$PCPD_{6\perp 3}$	0.13***	0.12***	0.13***	0.12***	0.12***	0.0852***
	(5.0099)	(5.0012)	(4.8652)	(4.5067)	(4.2522)	(3.7247)
$SBAS$	-0.0450	-0.0217				
	(-1.01)	(-0.5177)				
$OLBAS$	-0.0310	0.0061				
	(-0.66668)	(0.15228)				
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes
R^2	9.58%	8.87%	9.72%	9.05%	9.17%	8.48%
# Obs.	273,487	273,487	231,972	231,972	273,330	273,330

Table 11: **Rejecting Volatility Effects.**
This table presents evidence to rule out volatility effects. Columns 1-2 control for the straddle ($STRAD$), defined as $(CIV + PIV)/2$. Columns 3-4 uses $\%PCPD = PCPD/STRAD$ as regressors. Columns 5-6 uses $\%PCPD$ and control for $STRAD$ at the same time. Columns 7-8 finally include both volatility and microstructure controls in one regression. The horizon timing effects and the lower-diagonal pattern in the coefficients survive all tests. R^2 of each regression is also reported on the last row. 10%/5%/1% significance levels are marked with */**/***.

	Control for Volatility with Straddle		Control for Volatility using $\%PCPD$		Both $\%PCPD$ and Straddle		Control for Both Microstructure and Volatility	
Dependent Var.	R_{1-3} (1)	R_{4-6} (2)	R_{1-3} (3)	R_{4-6} (4)	R_{1-3} (5)	R_{4-6} (6)	R_{1-3} (7)	R_{4-6} (8)
$PCPD_{3,ATM}$	0.20***	0.0648***					0.20***	0.0617***
	(6.3695)	(3.0418)					(6.3635)	(2.8594)
$PCPD_{6\perp 3}$	0.12***	0.11***					0.12***	0.11***
	(4.9128)	(4.2606)					(4.8862)	(4.2562)
$\%PCPD_{3,ATM}$			0.20***	0.0570***	0.19***	0.0592***		
			(7.6486)	(2.7291)	(6.8958)	(3.2695)		
$\%PCPD_{6\perp 3}$			0.0990***	0.0870***	0.0939***	0.0851***		
			(4.7617)	(4.4704)	(4.8127)	(4.0509)		
$STRAD_{3,ATM}$	-0.13	-0.0565			-0.16	-0.0671	-0.15	-0.0632
	(-0.67672)	(-0.30097)			(-0.83001)	(-0.35816)	(-0.761)	(-0.34489)
$SBAS$							-0.0584*	-0.0194
							(-1.8083)	(-0.6048)
$OBAS$							-0.0160	-0.0094
							(-0.31002)	(-0.22059)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Normalized Coeff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	10.35%	9.44%	9.06%	8.41%	10.25%	9.35%	10.68%	9.76%
# Obs.	273,487	273,487	273,487	273,487	273,487	273,487	273,487	273,487