

THE UNIVERSITY OF  
**SYDNEY**

Economics Working Paper Series

2020 - 04

## Inferring Financial Bubbles from Option Data

Robert A. Jarrow and Simon S. Kwok

June 2021

# Inferring Financial Bubbles from Option Data

Robert A. Jarrow  
Cornell University

Simon S. Kwok  
The University of Sydney\*

First Version: May 2020

This Version: June 2021

## Abstract

Financial bubbles arise when the underlying asset's market price deviates from its fundamental value. Unlike other bubble tests that use time series data and assume a reduced-form price process, we infer the existence of bubbles nonparametrically using option price data. Under no-arbitrage and acknowledging data constraints, we can partially identify asset price bubbles using a cross section of European option prices. In the empirical analysis, we obtain interval estimates of price bubbles embedded in the S&P 500 Index. The estimated index bubbles are then used to construct profitable momentum trading strategies that consistently outperform a buy-and-hold trading strategy.

Keywords: asset price bubble, fundamental value, risk-neutral probability measure, state price distribution, partial identification, nonparametric estimation

JEL codes: C14, C58, G12

## 1 Introduction

The existence of asset price bubbles has been alledged for centuries, starting with the Dutch tulip bulb mania of 1634 (see Jarrow, Protter, and Shimbo (2010) for a brief summary of early alleged

---

\*We thank Eric Ghysels, the co-editor, the anonymous referees, and the participants in 2019 Time Series and Forecasting Symposium, 2021 NASMES, 2021 IAAE Annual Conference, and the seminar at UNSW. Their constructive comments led to substantial improvement of the paper. Jarrow: Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, N.Y. 14853 and Kamakura Corporation, Honolulu, Hawaii 96815; Email: raj15@cornell.edu. Kwok: School of Economics, The University of Sydney, NSW 2006, Australia; Email: simon.kwok@sydney.edu.au.

bubbles). Despite the longevity of these allegations, bubbles' existence is still debated because the literature differs both (i) on the definition of asset price bubbles (see Protter (2013)), and (ii) given a particular definition, on the joint hypotheses needed on future prices, cash flows, interest rates, and risk premiums to compute a bubble's magnitude.

Some methodologies rely on specific asset price processes. For example, Jarrow, Kchia, and Protter (2011a,b) assume that the asset price process follows a simple diffusion model. To detect bubbles, they apply Feller's test for explosions, estimated using a kernel method. Shiryaev, Zhitlukhin, and Ziemba (2014, 2015) specify a more complex diffusion process (*aka* stochastic disorder model) in which the drift and volatility parameters jump at change points. The bubble tests developed by Phillips, Shi, and Yu (2015a,b) look for explosive dynamics in long time series of asset prices using variants of an augmented Dickey-Fuller test, which is implemented using an autoregressive framework.

All of these approaches, however, focus on using time series price data. By contrast, the purpose of this paper is to infer the existence of bubbles using option price data, based on the recently developed local-martingale theory of bubbles. This theory provides a characterization of bubbles under a refinement of the notion of no-arbitrage, called "no free lunch with vanishing risk" (NFLVR). The first Fundamental Theorem of Asset Pricing (Delbaen and Schachermayer (1994, 1998)) asserts that NFLVR is equivalent to the existence of an equivalent local martingale measure (ELMM). Given an ELMM, a price process exhibits a bubble if and only if it is a strict local martingale under this ELMM.<sup>1</sup> This theory generates a number of results that characterize the impact of price bubbles on related call and put options (See Protter (2013) for a comprehensive survey of this new approach). These results provide the foundations for the empirical analysis in this paper.

Unlike time series price based approaches, the cross section of option prices contains useful information about the state-price distribution (SPD) of the underlying asset which is related to the risk-neutral measure. Breeden and Litzenberger (1978) show that the partial derivative of a European option price with respect to the strike price identifies the risk-neutral distribution and density function. Nonparametric methods that recover the SPD from option data are abundant

---

<sup>1</sup>A strict local martingale is a local martingale that is not a martingale. In the empirical option pricing literature, one typically assumes the existence of an equivalent martingale measure, which directly rules out the presence of bubbles.

in the literature (e.g., Aït-Sahalia and Lo (1998), Aït-Sahalia and Duarte (2003), Yatchew and Härdle (2006), Fan and Mancini (2009), Kitsul and Wright (2013), Song and Xiu (2016), Lu and Qu (2018)). In principle, we may also deduce the fundamental asset value (defined as the mean of the SPD) and obtain an estimate of the asset price bubble.

However, several empirical challenges arise. In terms of data quality, options with extreme moneyness have noisy prices, although they are the most informative with respect to the tails of the SPD. A more serious problem is that option prices are only available over a bounded range of strike prices. Many empirical works implicitly treat the bounded strike range as the support of the SPD.<sup>2</sup> Since the true state-price distribution has unbounded support (e.g., over the positive reals), the two tails are truncated due to missing option data, thus leading to bias in the probability and moment estimates based on the truncated probability distributions. An exception is the recent work of Lu and Qu (2018) which allows probability masses to spread beyond the bounded strike range.<sup>3</sup> An imposed parametric functional form may resolve the missing observation issue (e.g., Shimko (1993), Figlewski (2010)), but the modeler is still subject to the “curse of a joint hypothesis problem.”

In this paper, we do not impose any parametric model structure on the SPD, and at the same time acknowledge that the fundamental asset value (and hence the bubbles) cannot be point-identified due to missing observations in the tails. We show that it is possible to achieve partial identification of asset price bubbles under mild restrictions, namely NFLVR and the independence of call price bubbles on strike prices. The latter restriction is much weaker than the assumption of no-dominance, but it still facilitates the nonparametric identification of asset price bubbles using call prices in the presence of call price bubbles. We propose two methods to infer asset price bubbles: (i) a set of simple bounds expressed in terms of option prices with extreme moneyness, and (ii) a more refined nonparametric interval estimator that utilizes the entire cross section of option prices. Both methods do not assume no-dominance, which allows violations of put-call parity in market

---

<sup>2</sup>Nonparametric local estimation, which is popular in the literature, is silent about the tail parts of the distribution. Nonetheless, there are strategies to enlarge the strike range in practice, e.g., focusing only on options with a short time-to-maturity, pooling data across times, or using both calls and puts.

<sup>3</sup>Lu and Qu (2018) show that the risk-neutral density can be identified up to a set of constraints on the tail probabilities and moments. Their sieve estimation, however, is not suitable for inferring asset price bubbles because the sieve estimates using call data as an input would be contaminated by the existence of call price bubbles. Local polynomial fitting, which yields the estimate of the first-order derivative of the call price function, resolves this problem. See Section 3.

prices.<sup>4</sup>

We apply our bubble inference procedure to S&P 500 index time series price data. We first illustrate that the massive index options' data on any single day contains sufficient information to infer the size of S&P 500 bubbles on that day. With the combined use of call and put data, the nonparametric interval estimator indicates that bubbles are present and that they closely track market performance. In the second application, we construct momentum trading strategies based on the S&P 500 bubble estimates. The main finding is that profitable bubble-based momentum trading strategies exist, and they can earn consistently higher annual returns than the simple buy-and-hold and pure momentum strategies.<sup>5</sup> For the sample period spanning 1996-2015, the bubble-based momentum strategy (based on S&P 500 or its bubble deduced from call and put options expiring in three months) outperforms a pure momentum strategy by 1.8 percentage points and the buy-and-hold strategy by 3.6 percentage points, after netting out transaction costs. The superior performance suggests that the bubble estimates derived from option price data contains forward-looking information which, if used judiciously, may generate abnormal profits by “riding on the bubble” (Conlon (2004)).

Recently, Fusai, Jarrow and Lamichhane (2020) also proposes a new, but different methodology to detect and quantify bubbles in asset prices using option prices. In contrast to our approach, they use a parametric model for the evolution of the asset price process that facilitates the estimation of the option's fundamental value. To obtain their estimates, they impose the joint assumptions of NFLVR and no-dominance. Consistent with our results, they document the existence of bubble episodes in the S&P 500 index over the 2014-2018 time period.

The rest of the paper is organized as follows. Section 2 reviews the local-martingale theory of bubbles. Section 3 develops the methods for inference of asset price bubbles. Section 4 presents the empirical findings on the S&P 500 index. Section 5 concludes. Technical details about the proposed methodology and all proofs of theoretical results are collected in the Appendix.

---

<sup>4</sup>There exist empirical studies that report violations of put-call parity in certain time periods (e.g., Battalios and Schultz (2011), Nishiotis and Rompolis (2019)). In light of this empirical evidence, it is desirable to develop an inference method for asset price bubble that does not depend on the validity of put-call parity.

<sup>5</sup>The existence of profitable bubble-based trading strategies does not violate the local martingale theory because these strategies involve risk and do not constitute arbitrage opportunities.

## 2 The Local-Martingale Theory of Bubbles

This section reviews the local martingale theory of bubbles and presents the arbitrage-free option valuation formulas used to test for the existence of bubbles. The local martingale theory of bubbles was developed in a sequence of papers by Loewenstein and Willard (2000a,b), Cox and Hobson (2005), Heston, Loewenstein, and Willard (2007), and Jarrow, Protter, and Shimbo (2010). This theory extends the characterization of bubbles from discrete time, infinite horizon models to finite or infinite horizon, continuous trading models. It has been shown that rational bubbles, consistent with no-arbitrage and equilibrium, can exist in such settings (see Jarrow (2015) for a review of this literature). This approach facilitates the use of continuous time stochastic process mathematics to help estimate and to test for the existence of asset price bubbles.

Let us describe the market setting (see Jarrow, Protter, and Shimbo (2010), and Jarrow and Larsson (2012) for details). We start with a filtered complete probability space  $(\Omega, \mathcal{F}, \mathbb{F}, P)$  where  $\Omega$  is the state space,  $\mathcal{F}$  is a  $\sigma$ -algebra,  $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$  is a filtration that satisfies the usual hypothesis (e.g., Protter (2005)), and  $T$  is the terminal date which is assumed finite. The probability measure  $P$  is the physical measure. We assume that the market is frictionless (no transaction costs and restrictions on trade; e.g., short-selling and borrowing are permitted), competitive (all investors are price-takers), and incomplete (there are more random sources than traded securities). Market incompleteness is required so that bubbles can start, die, and be reborn within the model's horizon.

For simplicity, we assume that the market contains a money account and a single risky asset. The value of a unit of money account at time  $t$  is given by  $R_t := \exp\left(\int_0^t r_s ds\right)$  where  $r_s$  is the default-free spot rate of interest at time  $s$ . We assume that  $r$  is  $\mathbb{F}$ -adapted and that  $\int_0^T |r_s| ds < \infty$   $P$ -a.s. The market price of a share of risky asset at time  $t$  is denoted  $S_t$ . The price is the ex-dividend price if the asset pays dividends. We assume that the price process  $\{S_t\}_{t \geq 0}$ , is a non-negative, càdlàg semimartingale with respect to  $\mathbb{F}$ .

To facilitate further discussion of the market structure, we define the notion of trading strategies. A *trading strategy* is an  $\mathbb{F}$ -adapted vector process  $H := \{H_t\}_{t \geq 0}$  where  $H_t := (\eta_t, \pi_t)$  denotes the portfolio consisting of  $\eta_t$  units of money account and  $\pi_t$  units of the risky asset at time  $t$ . Denote the time- $t$  value of a unit share in each of these securities as  $U := \{U_t\}_{t \geq 0}$  where  $U_t := (R_t, S_t)$ . The portfolio value at time  $t$  is given by  $V_t := \eta_t R_t + \pi_t S_t$ . The stochastic integral  $(H \cdot U)_t :=$

$\int_0^t \eta_u dR_u + \int_0^t \pi_u dS_u$ , with the convention  $(H \cdot U)_0 = 0$ , represents the total profits generated by the trading strategy  $H$  over the period  $(0, t]$  starting from a zero initial investment.

The trading strategy is said to be *self-financing* if  $V_t = (H \cdot U)_t$  for all  $t$ . This implies that the portfolio starts with zero initial investment and that there are no cash inflows or outflows from the portfolio over time. An  $\mathbb{F}$ -predictable and  $U$ -integrable process  $H$  is said to be a  $\mathbb{F}$ -*admissible* trading strategy if  $H$  is self-financing and does not lead to unbounded loss with positive probability; more precisely, there exists a real number  $a > 0$  such that  $(H \cdot U)_t \geq -a$  for all  $t \geq 0$  almost surely.

Now we are ready to discuss the key market attributes for the study of bubbles. We adopt the no-free-lunch-with-vanishing-risk condition due to Delbaen and Schachermayer (1994, 1998a). This is essentially an asymptotic analog of the standard no-arbitrage condition.

**Assumption NFLVR:** There exists no sequence  $f_n = (H^n \cdot U)_T$  where each  $\mathbb{F}$ -admissible trading strategy  $H^n$  is such that, as  $n \rightarrow \infty$ ,  $\|\max(-f_n, 0)\|_\infty \rightarrow 0$  and  $f_n \rightarrow f$  *P-a.s.* for some  $f \geq 0$  with  $P(f > 0) > 0$ .

By the Fundamental Theorem of Asset Pricing (Delbaen and Schachermayer (1994, 1998a)), NFLVR holds if and only if there exists a nonempty set  $\mathcal{M}_{loc}$  of equivalent local martingale measures (ELMM) under which the wealth process  $V_0 + (H \cdot U)_t$  is a local martingale. The existence of an ELMM allows for the computation of the fundamental price of an asset, defined as the conditional expectation of the risky asset's discounted cash flows, where the expectation is taken under an ELMM.

In an incomplete and NFLVR market, there is more than one such measure in  $\mathcal{M}_{loc}$ , so the fundamental price is not unique and depends on the ELMM chosen by the market. The existence of this non-empty set of ELMMs allows for the dynamic nature of asset price bubble birth, death, and rebirth (see Section 4.2 of Jarrow, Protter, and Shimbo (2010)).

Another market attribute that is often imposed is the assumption of no dominance. The condition was introduced by Merton (1973) in the option pricing context and is formalized for the study of bubble process by Jarrow, Protter, and Shimbo (2010). We say that the  $i^{th}$  asset in the  $U^i \in (U^0, U^1) := (R, S)$  for  $i = 0, 1$  is *undominated* on  $[0, T]$  if there exists no  $\mathbb{F}$ -admissible trading strategy  $H$  such that  $U_0^i + (H \cdot U)_T \geq U_T^i$  *a.s.* and  $P[U_0^i + (H \cdot U)_T > U_T^i] > 0$ . In other words, the no-dominance condition on an asset holds if no admissible trading strategy can yield more payoff

than a buy-and-hold strategy in the asset itself.

**Assumption ND:** All assets (both the money account and the risky asset) are undominated on  $[0, T]$ .

Assumption ND rules out bubbles in a complete market (Jarrow, Protter, and Shimbo (2006)) but can accommodate different types of bubbles in an incomplete market (Jarrow, Protter, and Shimbo (2010)). A natural implication of ND is that put-call parity holds for the market prices of European options. While NFLVR is an individual-trader assumption, ND is a market-wide condition that governs the collective actions of investors (Jarrow and Protter (2010)).<sup>6</sup>

Suppose the market selects the ELMM  $Q_t^* \in \mathcal{M}_{loc}$  at time  $t$ . Under this probability measure, *the fundamental price* of the risky asset is defined by

$$\nu_t^* := E_t^*[R_t^{-1}R_T S_T], \quad (1)$$

where  $E_t^*[\cdot]$  denotes the expectation taken under  $Q_t^*$  conditional on  $\mathcal{F}_t$ , and  $S_T$  corresponds to the liquidation value of the risky asset at the model's horizon (at finite time  $T$ ). This is the expected (risk-adjusted using the ELMM) discounted cash flow from liquidating the stock at the model's horizon. Alternatively stated, it is the value agents would pay if after purchase, they had to hold the stock until liquidation. This fundamental value is consistent with the standard definition of an asset's fundamental value used in the classic economics literature (see Jarrow (2015) for the documentation of this claim).

When studying options, however, we need to generalize this definition to identify the fundamental value of the risky asset when the risky asset is sold *before time  $T$* . The reason for this generalization is that we are going to infer an asset's price bubble using option data. And, an option's payoff is defined relative to the market price of the risky asset on the option's maturity date, which is before  $T$ . Hence, the fundamental value of an option must be defined relative to the market price of the risky asset on the option's maturity date, and not the fundamental value of the risky asset on the option's maturity date as given in expression (1) above.

Hence, we define the time- $t$  fundamental price of the risky asset to be sold at time  $t + \tau$  (where

---

<sup>6</sup>We will not assume ND for bubble identification except for the naïve bound in Corollary 1.



$\tau \leq T - t$ ) to be

$$\mu_t^* := E_t^*[R_t^{-1}R_{t+\tau}S_{t+\tau}]. \quad (2)$$

The difference between expression (1) and (2) is clarified by rewriting (1) as

$$\nu_t^* = E_t^*[R_t^{-1}R_{t+\tau}\nu_{t+\tau}^*]$$

where

$$\nu_{t+\tau}^* = E_{t+\tau}^*[R_{t+\tau}^{-1}R_T S_T].$$

We see here that the generalized definition replaces the fundamental value of the risky asset at the selling time  $t + \tau$ ,  $\nu_{t+\tau}^*$ , with the market price of the risky asset,  $S_{t+\tau}$ .

The risky asset's *bubble associated with the selling time  $t + \tau$*  is defined as the difference between the market and fundamental prices of the asset:

$$\Pi_t^S(\tau) := S_t - \mu_t^*. \quad (3)$$

As defined, the bubble depends on the selling date of the asset. This is necessary because under the ELMM the asset price process is a supermartingale (Jarrow, Protter, and Shimbo (2010)). The expectation declines as the time to the asset's selling date decreases, hence, the dependence.

There is a related econometric literature that has explored bubble detection using historical time series asset prices and alternative methods (e.g., see Phillips, Shi, and Yu (2015a,b)). These papers use a different model structure and a different definition for an asset price bubble. Our model structure is continuous trading over a finite horizon; theirs is discrete trading over an infinite horizon. As discussed in Jarrow, Protter, and Shimbo (2010) and Jarrow (2015), in the most general model structure, there are three types of bubbles, called type 1, type 2, and type 3 bubbles. A type 1 bubble only exists in infinite horizon models, and it captures fiat money (a security with zero cash flows, but strictly positive value). A type 2 bubble also only exists in infinite horizon models and it corresponds to an asset whose price process (under the risk neutral probability measure) is a martingale but not a uniformly integrable martingale. Intuitively, the risk-adjusted expected discounted cash flows and liquidation value at time “infinity” do not equal the market price. Last,

a type 3 bubble only exists in continuous trading models and it corresponds to an asset whose price process is a local martingale, but not a martingale. In economic terms, the risk adjusted expected discounted cash flows and liquidation value at any finite time do not equal the market price. We study type 3 bubbles in our paper (expression (3) above), while the econometric literature studies type 2 bubbles.

Type 2 and 3 bubbles capture different economic phenomena. In an infinite horizon model, one is looking at long time horizons to do the estimation, and a bubble reflects the price process not converging to a terminal value. One can get explosive time series behavior. In a finite horizon model, a type 3 bubble corresponds to short horizon trading strategies that create value via buying and selling, independent of the fundamental value's cash flows (e.g. high frequency trading). Consequently, to test for type 2 bubbles, one needs a long-time period. By contrast, to test for type 3 bubbles, one can study short-time periods. This obvious sampling period difference is important in financial markets which exhibit structural shifts that invalidate the estimated model's assumptions. Over the past three decades, it is uncontested that our financial markets have undergone significant structural changes. Hence, testing for type 2 bubbles is problematic; but this is not true for type 3 bubbles.<sup>7</sup>

Let  $C_t(k, \tau)$  and  $P_t(k, \tau)$  be the time- $t$  market prices of the European call and put options<sup>8</sup> written on the risky asset with strike price  $k$  and time-to-maturity  $\tau$  (where  $\tau \leq T - t$ ). Under the ELMM, their time- $t$  fundamental prices are defined to be  $C_t^*(k, \tau) := E_t^*[R_t^{-1}R_{t+\tau}(S_{t+\tau} - k)_+]$  and  $P_t^*(k, \tau) := E_t^*[R_t^{-1}R_{t+\tau}(k - S_{t+\tau})_+]$ . Note that the options have an implied selling date for the underlying asset and the market price, not the fundamental value of the underlying asset, is in the payoff.

We define the options' bubbles to be the difference between their market and fundamental

---

<sup>7</sup>For type 3 bubbles, in a complete market, there is an equivalent definition of a bubble to that given above, which is that the asset's price process is dominated (see Heston, Loewenstein, and Willard (2000)). Our definition extends this complete market definition to an incomplete market where the market may satisfy ND. We point out that in a complete market, under NFLVR and ND, there are no bubbles.

<sup>8</sup>We will focus on European options throughout this paper. Our bubble inference results remain valid when American options are used, as long as the underlying asset pays no dividends and the jump process of the asset price satisfies certain regularity conditions (see Theorem 6.7 of Jarrow, Protter, and Shimbo (2010)).

prices, i.e.,

$$\Pi_t^C(k, \tau) := C_t(k, \tau) - C_t^*(k, \tau), \quad (4)$$

$$\Pi_t^P(k, \tau) := P_t(k, \tau) - P_t^*(k, \tau). \quad (5)$$

Jarrow, Protter, and Shimbo (2010) obtains some key results on the bubbles of risky assets in an incomplete market. In particular, under NFLVR, all put bubbles are identically zero, while call bubbles, like risky asset's bubble, are nonnegative in general. This holds regardless of Assumption ND. This is stated in the following proposition.

**Proposition 1** *Under Assumption NFLVR,  $\Pi_t^S(\tau) \geq 0$ ,  $\Pi_t^C(k, \tau) \geq 0$  and  $\Pi_t^P(k, \tau) \equiv 0$ .*

To explain the above result intuitively, we note that a bubble can only exist when an asset is purchased to sell in the future at a higher price. But, if the price of the asset (or option) is bounded above, then at some future time, this can no longer happen. By backward induction, rational traders will realize this possibility at earlier times as well, and the motive for this purchase disappears at all earlier times. Hence, assets whose price processes are bounded above can have no bubbles. This is the case for a put option.<sup>9</sup> By contrast, risky assets and call options can exhibit bubbles because their payoffs are unbounded above.

If in addition ND holds, then put-call parity holds for market prices, and it follows that  $\Pi_t^C(k, \tau) = \Pi_t^S(\tau) + \Pi_t^P(k, \tau)$  (Jarrow et al. (2010)). By above proposition, we deduce that  $\Pi_t^C(k, \tau) = \Pi_t^S(\tau) \geq 0$  for all  $k$ , and hence the call bubble is independent of the strike price.

The independence property of call bubble on strike is generally weaker than ND and does not imply put-call parity for market prices. When the latter fails, a call bubble can remain invariant to the strike price, and can be different from the asset's price bubble. This observation will play an important role in asset price bubble identification without ND and without put-call parity for market prices.

**Assumption INK:**  $\Pi_t^C(k, \tau) = \Pi_t^C(\tau)$  for all  $k > 0$ .<sup>10</sup>

<sup>9</sup>In practice, traded put prices may reflect "panic sales" during crises where puts with lower strike prices increase in value. This tends to decrease the estimate of the fundamental price (as there is more probability at the left tail of  $Q_t^*$ ). According to Proposition 1, although panic sales may change the fundamental value of the stock, the put's market price remains equal to its fundamental value under no-arbitrage.

<sup>10</sup>We rule out  $k = 0$  in the condition because the European call with zero strike is essentially the same as the

This bubble independence assumption is a mild restriction to impose on call option market prices. It can be shown that Assumption INK implies that, for all  $t$  and  $\tau$ , the market prices of calls satisfy  $C_t(k_1, \tau) > C_t(k_2, \tau)$  if  $k_1 < k_2$ .<sup>11</sup> This condition is likely to hold in option markets because its violation would immediately invoke market transactions.

In addition, we may characterize the time series dynamics of the European call bubble process  $\{\Pi_t^C\}_{0 \leq t \leq T}$ . In a static market ( $Q_t^*$  is time-invariant), the bubble process, if not trivially zero for all time  $t$ , is a non-negative strict local martingale (i.e., a local martingale that is not a martingale), and once it collapses it stays at level 0 until time  $T$ . By contrast, a dynamic market environment ( $Q_t^*$  is time-varying) allows for multiple bubble births and collapses through a regime-switching mechanism, and  $\{\Pi_t^C\}_{0 \leq t \leq T}$  is generally a non-negative local martingale. See Jarrow et al. (2010) for details. We adopt a dynamic market setting for the rest of our analysis.

### 3 Inferring Asset Price Bubbles from Option Data

Our goal is to infer the asset price bubble  $\Pi_t^S(\tau)$  of a risky asset without parametric assumptions. We first discuss the possibility of using option price to uncover the state-price distribution (SPD) and its mean – the fundamental price. We then introduce several methods that yield a feasible bound or interval estimate of  $\Pi_t^S(\tau)$  (Sections 3.2-3.3).

#### 3.1 Identifying the SPD and Fundamental Price from Option Data

Suppose a share of risky asset price is priced at  $S_t$  at time  $t$ . Let  $Q_t^*$  denote the ELMM chosen by the market at time  $t$ . With a slight abuse of notation<sup>12</sup>, we use the same notation  $Q_t^*$  for the conditional cumulative distribution function (cdf) of  $S_{t+\tau}$  (aka the state price at time  $t + \tau$ ) given  $\mathcal{F}_t$ . That is, the *conditional state-price distribution* is given by  $Q_t^*(s) := Q_t^*\{S_{t+\tau} \leq s\} :=$

---

underlying risky asset.

<sup>11</sup> Assume INK. Fix  $t$  and  $\tau$ . Drop these arguments from the notation. Assume that  $C(k_1) \leq C(k_2)$  for  $k_1 < k_2$ . Then,  $\Pi^C(k_1) + C^*(k_1) \leq \Pi^C(k_2) + C^*(k_2)$ , or,  $C^*(k_1) - C^*(k_2) \leq \Pi^C(k_2) - \Pi^C(k_1)$ . But  $C^*(k_1) > C^*(k_2)$  by the property of the fundamental call price. This implies that  $0 < C^*(k_1) - C^*(k_2)$ , and hence  $0 < \Pi^C(k_2) - \Pi^C(k_1)$ . This gives the contradiction.

The converse is not quite true. Suppose  $C(k_1) > C(k_2)$  for  $k_1 < k_2$ . It follows that  $\Pi^C(k_1) - \Pi^C(k_2) > C^*(k_2) - C^*(k_1)$ , where  $C^*(k_2) - C^*(k_1)$  is a negative number. It is therefore possible that  $\Pi^C(k_1) = \Pi^C(k_2)$ .

<sup>12</sup> More precisely, the conditional cdf depends on both the current time and the holding period, and hence should be indexed by both  $t$  and  $\tau$ . Since the holding period is kept fixed in the subsequent analysis, dropping  $\tau$  should not cause confusion.

$Q^*\{S_{t+\tau} \leq s | \mathcal{F}_t\}$ . Assume (without a slight loss of generality) a constant risk-free spot interest rate  $r$ , so that  $R_t = e^{rt}$ .<sup>13</sup>

Fix a given time  $t$  and time-to-maturity  $\tau$ . To reduce the notational burden, we suppress the time subscript  $t$  (e.g., in the ELMM and the conditional cdf, fundamental prices, bubbles and boundary strike prices) and the time-to-maturity argument  $\tau$  (e.g., in option price functions and bubbles) wherever necessary.

The following lemma, due to Breeden and Litzenberger (1978), is important for recovering  $Q^*$  as it provides the link between  $Q^*$  and the price function of European options.

**Lemma 1** *Under Assumption NFLVR, we have*

$$\frac{\partial C(k)}{\partial k} = -e^{-r\tau}[1 - Q^*(k)] + \frac{\partial \Pi^C(k)}{\partial k}, \quad (6)$$

$$\frac{\partial P(k)}{\partial k} = e^{-r\tau}Q^*(k). \quad (7)$$

Lemma 1 is fundamentally important for identifying  $\mu^*$  and hence asset price bubbles. Proposition 1 implies that bubbles are included in call prices but not in put prices, and so that we can obtain  $Q^*$  from put prices. Nonetheless, when call bubbles are independent of the strike (Assumption INK), they can be eliminated by taking the first-order partial derivative of the call price with respect to strike (as  $\frac{\partial \Pi^C(k)}{\partial k} \equiv 0$ ); consequently,  $Q^*$  can be recovered using *both* put and call price data.

How can we obtain the fundamental price from the state-price distribution? Based on the definition in (2), it is evaluated as follows:

$$\mu^* := \int_0^\infty e^{-r\tau} s dQ^*(s) = \int_0^\infty e^{-r\tau} [1 - Q^*(s)] ds, \quad (8)$$

where the last step follows from an integration by parts.

The fundamental price can be computed from (8) if  $Q^*$  is precisely known. In practice, however,  $Q^*$  is not revealed beyond the strike price range. Nonetheless, when evaluating  $\mu^*$ , we can obtain reasonable bounds on the contribution from the truncated tails of  $Q^*$ . This is made possible by the following result.

---

<sup>13</sup>The result directly generalizes to the case with time-varying and deterministic interest rate.

**Lemma 2** For finite  $\ell, u > 0$ , we have

$$\int_0^\ell e^{-r\tau}[1 - Q^*(s)]ds = e^{-r\tau}\ell - P^*(\ell),$$

$$\int_u^\infty e^{-r\tau}[1 - Q^*(s)]ds = C^*(u).$$

Applying the above lemma with  $u = u_c$  (the maximum strike for calls) and  $\ell = \ell_p$  (the minimum strike for puts), we can decompose  $\mu_t^*$  in (8) as follows.

$$\mu^* = \int_{\ell_p}^{u_c} e^{-r\tau}[1 - Q^*(s)]ds + [e^{-r\tau}\ell_p + C^*(u_c) - P^*(\ell_p)]. \quad (9)$$

The integral on the right side of (9) is the contribution to  $\mu^*$  from the middle portion of  $Q^*$ . The integration is over the available range of strike prices of option data and hence can be evaluated. The last term (in square bracket) is the contribution to  $\mu^*$  from the two tails of  $Q^*$ . It involves the fundamental prices of options  $C^*(u_c)$  and  $P^*(\ell_p)$ , which are unobserved but can be bounded. Indeed, it follows from Proposition 1 that

$$-P(\ell_p) \leq C^*(u_c) - P^*(\ell_p) \leq C(u_c). \quad (10)$$

The bounded range can be quite narrow in practice, as  $C(u_c)$  and  $P(\ell_p)$  are the prices of the most out-of-the-money (OTM) calls and puts, which are the most inexpensive among the cross section of options. This will be useful for constructing bounds on price bubbles.

### 3.2 Simple Bounds on Asset Price Bubbles

Our first method constructs simple bounds on asset price bubbles by pooling together both call and put price data when possible. A complication arising from this data is that the strike range of calls and puts are bounded and unlikely identical. To allow for this feature, we let  $[\ell_c, u_c]$  ( $[\ell_p, u_p]$ ) denote the range of strike prices for calls (puts), and we consider all (six) combinations of strike price ranges.

The following proposition provides the bounds on asset price bubbles. Assumption INK is crucial for identifying such bounds using call prices which contain bubbles.

**Proposition 2** *Suppose Assumptions NFLVR and INK hold. Define:*

$$\begin{aligned}
L_1 &= S_t - e^{-r\tau} \ell_c + P(\ell_c) - P(\ell_p) - C(\ell_c), & U_1 &= S - e^{-r\tau} \ell_c + P(\ell_c) + C(u_c) - C(\ell_c), \\
L_2 &= S - e^{-r\tau} u_p + P(u_p) - P(\ell_p) - C(u_p), & U_2 &= S - e^{-r\tau} u_p + P(u_p) + C(u_c) - C(u_p), \\
L_3 &= S - e^{-r\tau} \ell_p - C(\ell_p), & U_3 &= S - e^{-r\tau} \ell_p + P(\ell_p) + C(u_c) - C(\ell_p), \\
L_4 &= S - e^{-r\tau} u_c + P(u_c) - P(\ell_p) - C(u_c), & U_4 &= S - e^{-r\tau} u_c + P(u_c), \\
L_5 &= S - e^{-r\tau} \ell_c + P(u_p) - P(\ell_p) - C(\ell_c), & U_5 &= S - e^{-r\tau} u_p + P(u_p) + C(u_c) - C(\ell_c), \\
L_6 &= S - e^{-r\tau} \ell_p - C(u_c), & U_6 &= S - e^{-r\tau} u_c + P(\ell_p).
\end{aligned}$$

For fixed time  $t$  and time-to-maturity  $\tau$ , and for any  $\omega \in [0, 1]$ , we have:

- (a)  $\max\{0, \omega L_1 + (1 - \omega)L_2\} \leq \Pi^S \leq \omega U_1 + (1 - \omega)U_2$  for  $\ell_p \leq \ell_c < u_p \leq u_c$ ,
- (b)  $\max\{0, \omega L_3 + (1 - \omega)L_4\} \leq \Pi^S \leq \omega U_3 + (1 - \omega)U_4$  for  $\ell_c \leq \ell_p < u_c \leq u_p$ ,
- (c)  $\max\{0, \omega L_1 + (1 - \omega)L_4\} \leq \Pi^S \leq \omega U_1 + (1 - \omega)U_4$  for  $\ell_p \leq \ell_c < u_c \leq u_p$ ,
- (d)  $\max\{0, \omega L_3 + (1 - \omega)L_2\} \leq \Pi^S \leq \omega U_3 + (1 - \omega)U_2$  for  $\ell_c \leq \ell_p < u_p \leq u_c$ ,
- (e)  $\max\{0, L_5\} \leq \Pi^S \leq U_5$  for  $\ell_p < u_p < \ell_c < u_c$ , and
- (f)  $\max\{0, L_6\} \leq \Pi^S \leq U_6$  for  $\ell_c < u_c < \ell_p < u_p$ .

*Remark:* The bounds are generated by repeated applications of Lemmas 1 and 2. Wherever the strike ranges of calls and puts overlap, the bounds may be expressed in terms of the call or put prices, or a linear combinations of them (the weight  $\omega \in [0, 1]$  indicates the proportion of the bound expression contributed from calls).

As discussed in Section 3.2, the information about  $Q^*$  is not fully revealed beyond the bounded range of strike prices. This is why the asset price bubble is identified only up to an interval. The width of the interval indicates the degree of uncertainty of the bubble size. For cases (a)-(d), the maximum length of the interval is  $C(u_c) + P(\ell_p)$ , the sum of the most OTM call and put option prices (see (10)). The intervals for cases (e)-(f) are wider as the available call and put prices do not seamlessly cover the whole strike price range  $[\ell_p, u_c]$  (hence additional uncertainty is involved).<sup>14</sup>

<sup>14</sup>In our empirical application, the proportions of each of the six cases are as follows: (a) 90.6%, (b) 0.14%, (c) 4.9%, (d) 2.0%, (e) 2.4%, (f) 0.0%.

The following corollary gives much simpler bounds for  $\Pi^S(\tau)$  when the stronger Assumption ND is imposed.

**Corollary 1** *Under Assumptions NFLVR and ND, we have*

$$0 \leq \Pi^S(\tau) \leq C(u_c, \tau). \quad (11)$$

*Remark:* The bounds in (11) are sharper but less general than in Proposition 2 (where ND may fail). Indeed, the bounds in cases (a)-(d) of Proposition 2 reduce to (11) in the corollary when put-parity holds for market prices (a consequence of ND).

Intuitively, under ND, the asset price bubble is equal to the call bubble (see the discussion after Proposition 1), which is bounded from above by the call price, i.e.,  $\Pi^S = \Pi^C(k) = C(k) - C^*(k) \leq C(k)$  for any  $k$ . The least upper bound is obtained by selecting the most OTM call option, the cheapest one available in the cross section.

### 3.3 Interval Estimation of Asset Price Bubbles

The simple bounds in Proposition 2 (and Corollary 2) involve the prices of options with extreme moneyness. These options are usually thinly traded and subject to the most pricing uncertainty due to the lack of liquidity. The illiquidity of these options might affect the precision of the bound. This is also part of the reason (apart from short-sale constraints, recording error, etc.) that option prices are sometimes inconsistent with no arbitrage. On the other hand, options which have less extreme moneyness are more frequently traded. Since the transactions of these options are abundant in the data, they are informative about the middle part of  $Q^*$ . The unintended tail truncation is an artifact of missing observations of option prices beyond the strike price range, and will generally lead to bias when estimating asset price bubbles. Any practical methodology that attempts to infer the size of asset price bubbles needs to take into account the bounded and discrete nature of strike prices in option data. Our next proposed inference procedure delivers interval estimates of asset price bubbles that respect the aforementioned data constraints (e.g., discrete number of strike prices in a bounded range, inaccuracy of option prices) and are consistent with no arbitrage.



We introduce the tail-truncated state-price distribution over the support  $[\ell, u]$ , where  $\ell := \min(\ell_c, \ell_p)$  and  $u := \max(u_c, u_p)$ .

$$Q^\dagger(s) := \frac{Q^*(s) - Q^*(\ell)}{Q^*(u) - Q^*(\ell)} \quad \text{for } s \in [\ell, u]; \quad (12)$$

$Q^\dagger(s) = 0$  for  $s < \ell$ ; and  $Q^\dagger(s) = 1$  for  $s \geq u$ . This probability distribution is obtained by evenly redistributing probability mass in the tails of  $Q^*(s)$  to the strike price range  $[\ell, u]$ .<sup>15</sup> It is clear that  $Q^\dagger$  defines a valid probability measure with support  $[\ell, u]$ . Under this (incorrect) tail-truncated probability measure, the fundamental asset value is evaluated as  $\mu_t^\dagger := E_t^\dagger[e^{-r\tau} S_{t+\tau}]$  (where  $E_t^\dagger[\cdot]$  denotes the conditional expectation taken with respect to  $Q_t^\dagger$ ). It is given by (with subscript  $t$  suppressed)

$$\begin{aligned} \mu^\dagger &= e^{-r\tau} \left\{ \left( \int_0^\ell + \int_\ell^u + \int_u^\infty \right) [1 - Q^\dagger(s)] ds \right\} \\ &= e^{-r\tau} \left\{ \ell + \int_\ell^u [1 - Q^\dagger(s)] ds \right\}. \end{aligned} \quad (13)$$

The estimator  $\hat{\mu}^\dagger$  based on the tail-truncated distribution is biased for the fundamental asset value  $\mu^*$ , but, as we will discuss below, it is possible to correct and bound the bias.

Let us outline the steps for obtaining the point and interval estimates of the asset price bubbles (more details are found in the Appendix). The first step is to validate the option price data for further analysis. There may exist anomalies in the collected option data; in particular, the actual recorded prices may violate the theoretical relationships of option prices implied by no-arbitrage.<sup>16</sup> To mitigate this data issue, we follow Ait-Sahalia and Duarte (2003) by conducting a constrained least squares optimization on each cross section of option prices (over strike prices  $k$ , for fixed time  $t$  and maturity  $\tau$ ).<sup>17</sup> This step is detailed in Appendix A1. The constrained optimization yields a

<sup>15</sup>An alternative transformation is to use the estimator  $\check{Q}^*(s) = Q^*(s)$  for  $\ell_p \leq s < u_c$ ;  $\check{Q}^*(s) = 0$  for  $s < \ell_p$ ; and  $\check{Q}^*(s) = 1$  for  $s \geq u_c$ . This essentially lumps all the left tail mass to  $\ell_p$  and right tail mass to  $u_p$ . Since there is no natural *a priori* reason to load the tail probability mass to the boundaries, we do not pursue this transformation in the empirical study.

<sup>16</sup>By the no-arbitrage theory of option pricing (Black and Scholes (1973), Merton (1973)), the prices of European calls and puts must obey certain properties. For example, the traded call (put) price decreases (increases) with strike price (monotonicity) in a smooth way (convexity); e.g., see Theorem 8.5 of Merton (1973). Nonetheless, one can occasionally spot violations of these properties in actual option data. Some possible causes are recording error and illiquid trading. The imperfection of option data is at odds with the no-arbitrage theory, although it may not constitute an arbitrage opportunity in practice after taking into account market frictions (e.g., bid-ask spread and transaction costs), which the theory assumes away.

<sup>17</sup>The aim of Ait-Sahalia and Duarte (2003) is to recover the state-price density from option data. By contrast,

collection of adjusted option price data. We henceforth work with the adjusted option price data  $\{k_i, m_i\}_{i=1}^n$ .

Next, using the adjusted option data, we recover the state price distribution  $Q^*(\cdot)$  and the fundamental value. To this end, we estimate the call (put) price function  $m_c(\cdot)$  ( $m_p(\cdot)$ ) as a function of strike price, by fitting a local polynomial to the cross section of option prices. The fitting is done separately for each cross section of puts or calls and for each  $\tau$  and  $t$ . The estimator  $\hat{Q}_o^*(\cdot)$  ( $o = c, p$ ) of  $Q^*(\cdot)$  is constructed using the first-order derivative estimator  $\hat{m}_o^{(1)}(\cdot)$  of the call or put price function (by Lemma 1).<sup>18</sup> Define  $\hat{Q}^*(s) := \omega \hat{Q}_c^*(s) + (1 - \omega) \hat{Q}_p^*(s)$ , where  $\omega \in [0, 1]$  reflects the relative importance of call data in the estimation. Due to tail truncation, only those values of  $\hat{Q}^*(s)$  over  $s \in [\ell, u]$  can be obtained. To make it a proper distribution function, we renormalize  $\hat{Q}_o^*(s)$  ( $o = c, p$ ) to  $\hat{Q}_o^\dagger(s)$  according to (12). Define the distribution function  $\hat{Q}^\dagger(s) := \omega \hat{Q}_c^\dagger(s) + (1 - \omega) \hat{Q}_p^\dagger(s)$  to be the estimator of the tail-truncated cdf  $Q^\dagger(s)$ . The construction of  $\hat{Q}^\dagger(s)$  is summarized as follows ( $1\{o = c\}$  takes the value of one if call data are used, and is zero otherwise):

$$\hat{Q}_o^*(s) := e^{r\tau} \hat{m}_o^{(1)}(s) + 1\{o = c\} \quad \text{for } o = c, p, \quad (14)$$

$$\hat{Q}_o^\dagger(s) := \frac{\hat{Q}_o^*(s) - \hat{Q}_o^*(\ell_o)}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)} \quad \text{for } o = c, p, \quad (15)$$

$$\hat{Q}^\dagger(s) := \omega \hat{Q}_c^\dagger(s) + (1 - \omega) \hat{Q}_p^\dagger(s). \quad (16)$$

An estimator  $\hat{\mu}^\dagger$  of the fundamental value computed (incorrectly) under  $\hat{Q}^\dagger$  is then obtained by discretizing (13) over state prices  $s$  and substituting  $\hat{Q}^\dagger(s)$  for  $Q^\dagger(s)$ . More precisely, after choosing a grid of deterministic values  $\{s_i\}_{i=1}^{n_s}$  over  $[\ell, u]$  with  $s_0 = \ell$ ,  $s_{n_s} = u$ , and step sizes  $\Delta s_i := s_i - s_{i-1}$ , we evaluate  $\hat{\mu}^\dagger$  using the following relations:

$$\hat{\mu}^\dagger := e^{-r\tau} \ell + e^{-r\tau} \sum_{i=1}^{n_s} [1 - \hat{Q}^\dagger(s_i)] \Delta s_i. \quad (17)$$

The asymptotic normality of the local polynomial estimator is well known (Fan and Yao (2005), Theorem 6.3) and will be useful for the large-sample inference of  $\mu^\dagger$ .

---

our target parameter is the mean of the state-price distribution, which can be estimated as a function of the cdf as in (8). Compared to density estimation which requires the second-order derivative, it is more efficient to obtain the cdf estimate involving only the first-order derivative.

<sup>18</sup>Motivated by bias and variance considerations, Fan and Gijbels (1996) recommend the use of a local quadratic fit for estimating the first-order derivatives. See Appendix A1 for details.

Due to unintended tail truncation, there is a bias in  $\hat{\mu}^\dagger$  as an estimator for the fundamental value  $\mu^*$ . The following result is useful for computing the bias.

**Lemma 3** *The difference  $\mu^\dagger - \mu^*$  takes the form*

$$\mu^\dagger - \mu^* = B - A,$$

where  $A = C^*(u) - P^*(\ell)$  and

$$B = \frac{e^{-r\tau} Q^*(\ell)(u - \ell)}{Q^*(u) - Q^*(\ell)} - \left( \frac{1}{Q^*(u) - Q^*(\ell)} - 1 \right) e^{-r\tau} \int_\ell^u Q^*(s) ds.$$

*Remark:* The first term  $B$  in the difference  $\mu^\dagger - \mu^*$  arises from tail truncation. It vanishes as the strike price interval expands on both sides, i.e.,  $Q^*(\ell) \rightarrow 0$  as  $\ell \rightarrow 0$  and  $Q^*(u) \rightarrow 1$  as  $u \rightarrow \infty$ . This term is point-identified and can be feasibly estimated under Assumptions NFLVR and INK (using Lemma 1). The second term  $A$ , which arises from missing observations beyond the strike range, is partially identified and is bounded by (10).

Let  $\hat{B}$  be the feasible estimator of  $B$  by replacing the unobserved  $Q^*(s)$  with the sample analog (see Appendix A1). Define  $\hat{\mu} := \hat{\mu}^\dagger - \hat{B}$  as the estimator of the fundamental asset value with the bias term  $B$  corrected. This leads to the bias-corrected point estimator of the asset price bubble:

$$\hat{\Pi}^S := S - \hat{\mu} = S - \hat{\mu}^\dagger + \hat{B}. \quad (18)$$

It is important to note that  $\hat{\Pi}^S$  is *not* consistent for the asset price bubble as the latter is not point-identified. Nonetheless, it is possible to obtain a consistent interval estimate of the asset price bubble. The interval estimate accounts for both the sample variation of  $\hat{\mu}^\dagger$  and the partial identification of  $\mu^*$  (due to term  $A$ ). The validity of the confidence interval using a normal approximation is justified by the following proposition. Let  $h$  be the bandwidth of the local polynomial fit. Recall that  $n$  is the size of the option sample in the cross section.

**Proposition 3** *Suppose Assumptions NFLVR and INK hold, and that regularity conditions (a)-(e) (stated in Appendix A2) are satisfied. Let  $\{s_i\}_{i=1}^{n_s}$  be a sequence of deterministic numbers in*

$[\ell, u]$  such that  $\Delta s_i = 1/n_s$  for all  $i$ , and that  $n_s$  satisfies  $\frac{\log n}{n_s} = O(1)$  as  $n, n_s \rightarrow \infty$ . Then, as  $n, n_s \rightarrow \infty$ ,  $h \rightarrow 0$  and  $\frac{\log(n)}{nh^3} \rightarrow 0$ , we have

$$\hat{\mu}^\dagger \xrightarrow{p} \mu^\dagger$$

and

$$\sqrt{nh^3}[\hat{\mu}^\dagger - \mu^\dagger - O(n_s^{-1} + h^2)] \xrightarrow{p} N(0, V),$$

where  $V := \text{plim } V_{n_s, n, h}$  and

$$V_{n_s, n, h} := (nh^3) \text{Var} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} \hat{m}^{(1)}(s_i) \right).$$

By the normal approximation, the  $100(1 - \alpha)\%$  confidence interval for the asset price bubble  $\Pi^S$  is given by:

$$[\hat{\Pi}^S - z_{\alpha/2} \cdot s.e.(\hat{\mu}) - A_u, \hat{\Pi}^S + z_{\alpha/2} \cdot s.e.(\hat{\mu}) - A_\ell], \quad (19)$$

where  $z_{\alpha/2}$  is the standard normal variate with right tail probability  $\alpha/2$ . The standard error  $s.e.(\hat{\mu})$  accounts for the sample variability in  $\hat{\mu}^\dagger$  and  $\hat{B}$ . It is computed by the Newey-West method which allows for heteroskedasticity and weak cross-sectional correlations of option prices across strikes. Details are found in the Appendix.

## 4 The Empirical Study

This section applies the previous theory to study the existence of price bubbles in the S&P 500 index using call and put option prices. The estimated bubbles are then used to construct profitable momentum trading strategies.

### 4.1 The Data

We collect daily prices of European call and put options written on the S&P 500 index spanning 1996 to 2015 from OptionMetrics. We filter the option data by retaining only those trade records with positive volume, best bid or offer prices at least \$0.05, a time-to-maturity longer than eight days and no longer than a year, and an expiry date before the 27<sup>th</sup> of a month. This last filter

effectively removes all “non-standard” option contracts.<sup>19</sup> To ensure sufficient observations for estimation, we remove those maturity groups that have fewer than three observations.<sup>20</sup> The above filtering criteria are largely in line with data-cleaning practices commonly adopted in the empirical option pricing literature. An exception from common practices is that we retain all the deep in- and out-of-the-money options, which are useful to bound the asset price bubbles as shown in Propositions 2 and 3. The option prices used for further analysis are the average of the bid and offer prices.

The risk-free interest rate is the one-month T-bill rate released by the Federal Reserve Board. If the one-month rate is unavailable (which may be the case in earlier years in the sample period), the three-month T-bill rate is used instead.

## 4.2 The Information Content of Option Prices

In this section we examine the information content contained in options written on the S&P 500 index. Lemma 1 suggests that call and put index options provide information on the risk-neutral state-price distribution  $Q_t^*$  of the future realizations of the S&P 500 index. The theory, however, does not tell us whether we should use calls or puts or both of them to recover  $Q_t^*$ . Nor is it clear which of the bubble inference methods are more appropriate empirically. Some exploratory analysis on the option data provides guidance on the methods to use.

First, we examine the trading volume of the S&P 500 index options. Figure 1 plots the three-month moving average of the daily trading volume of different options (classified by option types and moneyness) over time (1996-2015). From the time series plot, we observe that, in general, more puts are traded than calls and that, for each option type, the ATM (at-the-money) options are the most heavily traded, followed by the OTM (out-of-the-money) and ITM (in-the-money) options. The popularity of the ATM put options increases sharply in 2007 when the S&P 500 index reached its peak before the global financial crisis.

Table 1 allows for a more refined comparison of the trading volume across option types, moneyness and maturity groups. Put options account for 63 percent of all the traded index options

---

<sup>19</sup>The “standard” options are those option contracts that expire on or before the third Friday of the expiration month.

<sup>20</sup>A *maturity group* is defined as the collection of options that share the same option type (call/put) and range of maturities. In our analysis, we classify options into three time-to-maturity groups with  $\tau$  spanning up to 3 months, 3-6 months, and 6-12 months (i.e.,  $\tau \in [0, 0.25]$ ,  $[0.25, 0.5]$  and  $[0.5, 1]$ ).

expiring within a year. Among the traded put (call) options in the sample, 83 (84) percent expire within three months. The OTM puts are nearly six times more actively traded than are the OTM calls. The average daily trading volume of the OTM puts is 63,600 contracts (which is 32 percent of all traded puts in the sample). This contrasts to a daily volume of 10,800 contracts for the OTM calls (which takes up less than 10 percent of all traded calls). OTM options are especially informative about the tails of the state-price distribution; in particular, OTM puts (calls) reflect the downside (upside) risk associated with the left tail (right tail). This vibrant trading activity for both calls and puts (especially ATM and OTM options) suggests that we would ignore much trading information by considering just the puts or calls as the data used.

We now study the range of strike prices. As discussed in Section 3.3, the span of strike prices crucially determines the revealed part of the SPD. Figure 2 documents the strike price ranges for all calls (green-colored band) and puts (blue-colored band) over time.<sup>21</sup> These strike prices represent the support of  $Q_t^*$  revealed by call and put prices. The overlapping region is the part of the support revealed by both the calls and puts. There is clearly some persistent asymmetry in the manner in which call and put prices recover  $Q_t^*$ . At any point in time, the cross section of all the put data reveals a much wider range of the discounted future state prices, especially those values at the lower end (downside) that are revealed only by OTM puts. By contrast, the call data, in particular OTM calls, tend to reveal the upper part (upside) of the SPD. With reference to the discussion in Section 3.3, the SPD is subject to less truncation when only puts are used for estimation (instead of calls). From a bias-reduction perspective, the bubble estimates derived using both calls and puts is preferred, followed by that derived from purely puts and purely calls.

Next, we investigate the validity of put-call parity. Define  $D_t(k, \tau) := S_t - [C_t(k, \tau) - P_t(k, \tau) + ke^{-r\tau}]$ , which is the discrepancy between the S&P 500 index and its “synthetic” version (in squared bracket) derived from the call and put prices with the same strike price and maturity. The discrepancy is equal to zero if put-call parity holds in market prices.

Figure 3 plots the discrepancy averaged over all matched strike prices and over each of the three maturity groups. We observe a generally non-zero discrepancy except in 2000 and near the end of 2008 when the actual index dips below the “synthetic” index. Nonetheless, the magnitude of the

---

<sup>21</sup>The three-month moving averages are plotted to smoothen the data. Also plotted in Figure 1 are the narrower ranges of strike prices (shown in dashed and dash-dotted lines) of ATM calls and puts, where ATM options are defined with absolute log-moneyness less than 0.1.

discrepancy is quite small relative to the index level (e.g., the average discrepancy is  $< 0.2$  percent of S&P 500 index for  $\tau \in [0, 0.25]$ ).

The discrepancy measure also reveals the extent to which empirically call bubbles depend on the strike price. Indeed, from the definition of option bubbles and by put-call parity in fundamental prices, it follows that  $D_t(k, \tau) = \Pi_t^S(\tau) - \Pi_t^C(k, \tau)$ . This provides a way to validate the empirical validity of strike independence of call bubbles (Assumption INK), which our proposed methods presume as an identification condition for asset price bubbles when call option prices are used. Figures A1-A3 show the same plots as in Figure 3, except that the discrepancy measure is averaged over strike prices associated with ITM, ATM and OTM calls (equivalently, OTM, ATM and ITM puts), respectively. Assumption INK is empirically confirmed if the plots in Figures A1-A3 look alike. The plots reveal that the discrepancy remains stable over the strikes associated with ITM and ATM calls (OTM and ATM puts) (Figures A1-A2), while the discrepancy associated with OTM calls (ITM puts) takes on somewhat smaller values (Figure A3). Since OTM calls and ITM puts have a much lower trading volume (see Figure 1 and Table 1), Assumption INK seems to be supported by the empirical evidence.

From the above empirical observations, it is desirable to adopt a method which does not depend on the validity of put-call parity and uses both call and put data. The strike price independence of call bubbles is the assumed structure which our proposed method employs to identify asset price bubbles.

### 4.3 Inferring the S&P 500 Bubble

In this section, we implement the methods in Section 3 to infer the presence of bubbles in the S&P 500 index using option prices. Recall that the bubble embedded in an option is a function of its time-to-maturity  $\tau$ . Since there are multiple strike options associated with different maturities, we classify the option data into three maturity groups: 0-3 months, 3-6 months, and 6-12 months (i.e.,  $\tau \in [0, 0.25]$ ,  $[0.25, 0.5]$  and  $[0.5, 1]$ , measured in years). We then compute the average bubble within each maturity group.

The naïve bounds (Corollary 1) assume the put-call-parity being valid. The upper bound, denoted  $\tilde{\Pi}_1(\tau)$ , is essentially the price of the most OTM call and is a function of the time-to-maturity.<sup>22</sup>

---

<sup>22</sup>For notational simplicity, we suppress the time subscript  $t$  in all bubble estimates in the empirical discussion

Figure 4 plots the moving averages of  $\tilde{\Pi}_1(\tau)$  associated with the three maturity groups.<sup>23</sup> We observe two major spikes: the first one occurs near the end of 1990's, another one at the end of 2007. The surge is particularly pronounced for  $\tau \in [0.5, 1]$ . These bubble bounds need to be interpreted with care in light of the thin trading volume of deep OTM call options and the empirical violation of put-call parity (see the end of Section 4.2).

The simple bounds (Proposition 2) do not rely on the validity of put-call-parity. We evaluate these bounds by setting  $\omega$  equal to the proportion of call options in the cross section. Figure 5 shows the moving averages of these bounds and their mid-points, denoted  $\tilde{\Pi}_2(\tau)$ , for each maturity group. We note that the width of the interval changes over time, reflecting the varying level of uncertainty of the bubble size. The interval tends to be wider during more volatile periods (e.g., mid-1997 to 2002, end of 2007 to 2008, and 2015). The interval also tends to be wider and more volatile for longer maturities. Similar to the naïve bounds, the precision of the simple bounds are subject to the illiquidity of options at extreme moneyness.

Nonparametric interval estimation (Section 3.3) delivers interval estimates of bubbles using the entire cross section of option prices. Like the simple bounds, it does not presume put-call parity. Figure 6 displays the moving averages of the point estimate (18) with  $\omega = n_c/n$  (point estimate denoted  $\hat{\Pi}_{cp}(\tau)$ ;  $n_c$  is the number of call contracts in the data), and the associated 95% confidence bounds (19). The confidence bounds keep track of the market performance more closely than the naïve and simple bounds in Figures 4 and 5. We observe that bubbles exist, although they are small most of the time.<sup>24</sup> Bubbles were positive and economically significant when the market was thriving in the first half of 1998, 2006-2007 and 2014-2015. Bubbles tend to be larger when the option's maturity (viewed as the time-to-liquidation of the S&P 500 portfolio) is larger. This is consistent with the supermartingale property of the bubble process (see the discussion after expression (3)).

The bubble estimates dip into the negative domain during 2009 (for  $\tau \in [0, 0.25]$ ) and near the end of 2011 (for  $\tau \in [0.25, 0.5]$ ). While the results are puzzling, the negative estimates occur when

---

(unless it is necessary to avoid causing confusion in the context).

<sup>23</sup>As the inference methods are designed for cross-sectional option data, the time series of bubble bounds/estimates change over time. To smooth the time series and facilitate the visual presentation, we plot the three-month moving average of all bubble estimates.

<sup>24</sup>One reason is that bubbles are generated when stocks are purchased to resell rather than to hold for a long time. These two are almost always equal, except in trading frenzies. When trading frenzies occur, volatility increases, and the size of the bubbles expands.



the market is low. This inconsistency with the local-martingale theory of bubbles can be explained by the facts that we are assuming both frictionless and competitive markets. When markets are falling, both of these assumptions are less likely to be satisfied.<sup>25</sup> Nonetheless, these violations are infrequent in our estimates.

Figures 7-8 show the analogous plots for  $\omega = 0$  (formed by puts only, point estimate denoted  $\hat{\Pi}_p(\tau)$ ) and  $\omega = 1$  (formed by calls only, point estimate denoted  $\hat{\Pi}_c(\tau)$ ) respectively. The bubble estimates are insignificantly different from zero over almost all time periods. Compared with  $\hat{\Pi}_c(\tau)$  and  $\hat{\Pi}_p(\tau)$ , it is clear that  $\hat{\Pi}_{cp}(\tau)$  has a much narrower confidence interval. This is due to both a reduction in sample variability for the local polynomial estimate (leading to smaller  $s.e.(\hat{\mu})$ ) and a widening of the strike price range (leading to sharper bounds on  $A$ ); see (19).

In summary, these empirical findings provide strong evidence for the existence of bubbles in the S&P 500 index. And, it provides equally strong evidence supporting the validity of the local-martingale theory of bubbles. Because the bubble estimates are derived from option data, one may wonder whether they capture forward-looking information about the future prospects of the market. Figure 9 displays the cross-correlograms between the bubble point estimates  $\hat{\Pi}_{cp}(\tau)$  and S&P 500 index, i.e.,  $Corr(\hat{\Pi}_{cp,t}(\tau), S_{t+m})$  against lag  $m$ , where one lag represents a separation of 10 days (two business weeks). The cross-correlograms reveal very persistent association from  $\hat{\Pi}_{cp}(\tau)$  to the S&P 500 index.<sup>26</sup> For the case  $\tau \in [0.25, 0.5]$  and  $[0.5, 1]$ , the cross-correlations are particularly strong and remain significant over multiple lags. This provides some indirect evidence suggesting that our bubble estimates predict future market performance. More direct evidence can be obtained by forming trading strategies that profit from its use. This is the content of the next section.

#### 4.4 Bubble-based Trading Strategies

To a financial economist, a bubble reflects the market's information about the fundamental value of an asset. Any discrepancy between the market and fundamental prices may indicate a dise-

---

<sup>25</sup>The presence of market frictions during crises can induce a liquidity differential between traded stocks and options. As seen in Figure 2, the S&P 500 falls below the lower bound of the strike price range of traded calls (in green) during the global financial crisis of 2008 and in the second half of 2011. However, the S&P 500 stays within the strike price range for put options (in blue). This indicates that the speed at which call option trading shifts towards the lower tail of the strike price range does not catch up with the rapid decrease of the S&P 500.

<sup>26</sup>The interesting results are the significant cross-correlations for positive lags. The cross-correlations for negative lags are expected as the bubble estimates are functions of option prices, which depend on past market performance.

equilibrium and trigger a series of adjustments in the market prices. Evaluating the out-of-sample performance of the estimated bubbles may provide insights into the price adjustment process and shed light on market inefficiency. In addition, an investor may take advantage of information on bubbles to design a profitable trading strategy (buy when the bubble is expanding and sell before it bursts). Such a strategy is motivated by the “greater fool” theory regarding the formation of bubbles (see Conlon (2004)). There exists some empirical evidence that trading strategies based on this principle can be profitable (Shiryaev, Zhitlukhin, and Ziemba (2014, 2015), Milunovich, Shi, and Tan (2019)).

To investigate the profitability of such strategies, we consider a joint momentum strategy based on the S&P 500 index (denoted  $S$ ) and its bubble estimate  $\hat{\Pi}^S(\tau)$ . More specifically, let  $MA_i(S)_t$  and  $MA_i(\hat{\Pi}^S(\tau))_t$  be, respectively, the  $i^{\text{th}}$  ( $i = 1, 2$ ) moving averages of  $S$  and  $\hat{\Pi}^S(\tau)$  over the window  $\{t - w_i + 1, t - w_i + 2, \dots, t\}$  of length  $w_i$ . The moving average windows are such that  $w_1 < w_2$ . The strategy suggests holding the S&P 500 index at time  $t$  if

$$MA_1(\hat{\Pi}^S(\tau))_t > MA_2(\hat{\Pi}^S(\tau))_t \quad \text{or} \quad MA_1(S)_t > MA_2(S)_t, \quad (20)$$

i.e., we take a long position when either the bubble estimate or the S&P 500 index (or both) exhibits an upward trajectory. If the investor does not hold the S&P 500 index, the proceeds are deposited into a money account with zero interest.

The bubble estimate is taken to be any one of the estimates/bounds studied in Section 3. This trading strategy is more aggressive than the pure momentum strategy based on S&P 500 itself (i.e., hold the market portfolio if  $MA_1(S)_t > MA_2(S)_t$ ). Another strategy is to buy-and-hold the S&P 500 index over the sample period. We investigate all three strategies. For our baseline results, the two moving averages are computed over a window of length  $w_1 = 3$  months and  $w_2 = 1$  year, and the portfolio is updated on a weekly basis.

The first three columns of Table 2 report the performance of different trading strategies over 1996-2015. The joint momentum strategy (20) with the bubble estimate chosen to be  $\hat{\Pi}_{cp}(\tau)$  or  $\hat{\Pi}_p(\tau)$  achieves superior performance: e.g., it generates an annual net return (after subtracting a round-trip 0.5% transaction cost) of 8.7 percent with bubble estimate  $\hat{\Pi}_{cp}(\tau)$  for  $\tau \in [0, 0.25]$ , whereas a pure momentum strategy based on  $S$  alone achieves 6.9 percent, and a buy-and-hold

strategy earns 5.1 percent. It also dominates the remaining joint momentum strategies constructed from the other bubble estimates. The conclusion is qualitatively similar after adjusting for risk, as revealed by the Sharpe ratios.

This additional net profit is attributable to the predictability of the bubble estimates derived from both put and call option prices. As shown in Figure 10, the joint momentum strategy takes advantage of the dynamics of  $\hat{\Pi}_{cp}(\tau)$ , which tends to send “buy” signals earlier than the market index does after a substantial market drop. The combined measure  $\hat{\Pi}_{cp}(\tau)$  seems to predict better the movement of the market index than the pure bubble estimates  $\hat{\Pi}_p(\tau)$  and  $\hat{\Pi}_c(\tau)$ . Indeed, compared to the joint momentum strategy based on  $S$  and  $\hat{\Pi}_{cp}(\tau)$ , the one based on  $S$  and  $\hat{\Pi}_p(\tau)$  generates slightly lower net returns (except for the maturity group  $\tau \in [0.25, 0.5]$ , for which the net return is marginally higher). The boost in the performance of bubble-based trading strategies seems to come from the combined use of calls and puts.

The rest of Table 2 shows the trading strategies’ performance over three sub-periods, each spanning five years. The performance of the joint momentum strategy based on  $S$  and  $\hat{\Pi}_{cp}(\tau)$  remains stable and robust for the shortest maturity group  $\tau \in [0, 0.25]$  – the strategy yields persistently high annual net return in all of the sub-periods, outperforming the benchmark of the pure momentum strategy on  $S$  and the buy-and-hold strategy. For  $\tau \in [0, 0.25]$ , the joint and pure momentum strategies perform equally well in 1996-2000, indicating that the buy signals are solely triggered by the S&P 500 index during this period. As the maturity becomes longer with  $\tau \in [0.5, 1]$ , the performance depends on the sub-periods: the strategy outperforms the benchmark of the pure momentum strategy on  $S$  by at least three percentage points per annum in 2006-2010 and 2011-2015, but becomes worse than the pure momentum strategy on  $S$  in 1996-2000 and 2001-2005. Replacing  $\hat{\Pi}_{cp}(\tau)$  with  $\hat{\Pi}_p(\tau)$  in the joint momentum strategy yields similar but slightly less stable performance, while using other bubble measures ( $\hat{\Pi}_c(\tau)$ ,  $\tilde{\Pi}_1(\tau)$  and  $\tilde{\Pi}_2(\tau)$ ) results in more variable (sometimes negative) net returns.

Next, we check the robustness of performance for different trading frequencies. The results are presented in Table 3. The joint momentum strategies that use  $\hat{\Pi}_{cp}(\tau)$  or  $\hat{\Pi}_p(\tau)$  outperform those adopting other choices of bubble estimates over all trading frequencies ranging from one to 10 days. The performance slightly deteriorates as the trading frequency decreases to 20 days. This indicates that both the bubble estimates and the index contain timely information about the changing market

sentiment and hence trading signals, which would be costly to the investors if slowly acted upon.

We also study the effect of changing the window lengths ( $w_1$  and  $w_2$ ) of the moving averages in the design of the momentum strategies. This may be regarded as a sensitivity analysis that calibrates the tuning parameters  $w_1$  and  $w_2$ , which are unknown to the investors *a priori*. Table 4 records the performance of trading strategies for various window lengths. The joint momentum strategies that adopt  $\hat{\Pi}_p(\tau)$  and  $\hat{\Pi}_{cp}(\tau)$  for  $\tau \in [0, 0.25]$  exhibit robustness in performance: they outperform the pure momentum strategy and the buy-and-hold strategy as  $w_1$  varies over 1-6 months and  $w_2$  is set to one year. The joint momentum strategies using other bubble estimates/bounds do not yield performance as robust as those of  $\hat{\Pi}_p(\tau)$  and  $\hat{\Pi}_{cp}(\tau)$ . The performance becomes more varied when using bubbles for longer maturities ( $\tau \in [0.25, 0.5]$  and  $\tau \in [0.5, 1]$ ) or when the long-span window varies in length ( $w_2$ ).

One may be concerned that the robust performance of the above trading strategies is a result of data snooping. To investigate their superior performance, while allowing for multiple testing, we compute the Bootstrap Reality Check test of White (2000). Here, the best performing trading strategy is tested against other similar trading strategies, but with different combination of trading frequencies (every 1, 5, 10, 20 days) and different lengths for the moving average window  $\{(w_1, w_2) : w_1 = 21, 42, 63, 84, 105, 126, w_2 = 126, 252, 378, 504, w_1 < w_2\}$  (for joint momentum strategies). The benchmark is selected to be the buy-and-hold strategy. To implement the stationary bootstrap, we consider a range of average block sizes spanning 3-12 months. Longer blocks tend to better preserve the serial dependence of the price and bubble processes – a dynamic property that momentum strategies rely on to generate profits. Table 5 reports the p-value of the Bootstrap Reality Check test. The tests confirms the superiority of the best performing joint momentum and pure momentum strategies against the buy-and-hold strategy. We note that the evidence is weaker when the average block size is three months.

Table 6 investigates the effect of redesigning the trading strategies. The baseline case of the joint momentum strategy defined in (20) (first pair of columns) is compared to momentum strategies purely based on the bubble estimates (second pair of columns), and to momentum strategies based on  $S$  and the bubble estimates. The holding rule (20) is modified by replacing “or” with “and.” These alternative strategies are less aggressive in taking a long position in the S&P 500 index, and thus yield substantially lower net returns than the baseline case.

All of the above results assume a round-trip transaction cost of 0.5 percent. As our final robustness check, we study the impact of transaction costs on trading performance. Table 7 reports the net returns of trading strategies for various round-trip transaction costs ranging from 0 percent (gross returns, first column), 0.5 percent (baseline case, second column) to 3 percent. The joint momentum strategy based on  $S$  and  $\hat{\Pi}_{cp}(\tau)$  with  $\tau \in [0, 0.25]$  becomes the only joint momentum strategy that uniformly outperforms the pure momentum strategy based on  $S$ . The uniform superiority of this trading strategy is not affected by the choice of the trading frequency.<sup>27</sup> Due to more frequent trading, other joint momentum strategies perform less well as the transaction cost increases. Using  $\tilde{\Pi}_1(\tau)$  or  $\tilde{\Pi}_2(\tau)$  to form the joint momentum strategy could be a decent alternative when transaction costs are low, but the need to trade frequently would lead to excessive transaction costs that eventually offset the positive returns.

## 5 Concluding Remarks

In this paper, we study a number of model-free asset price bubble inference methods that use option prices. Due to limited availability of option data (e.g., missing observations beyond a bounded range of strike prices), the state-price distribution is subject to tail truncation and cannot be fully recovered. We show that, under mild restrictions implied from the local-martingale theory of bubbles, the fundamental asset value and hence the asset price bubbles associated with a given maturity can be partially identified with a bounded range of available strike prices. To infer asset price bubbles from European option data, we propose two feasible methods: a general set of bounds in terms of option prices with extreme moneyness, and a more refined interval estimate that utilizes the entire cross section of both call and put data.

We illustrate the usefulness of the inference method in the empirical analysis of S&P 500 bubbles. The estimated bubbles track the market performance closely, and remain largely positive in the first half of 1998, during 2006-2007, and during 2014-2015. As an out-of-sample analysis, we show that it is possible to construct profitable momentum trading strategies based on both the S&P 500 index and the estimated bubbles. For our sample that spans 1996-2015, the joint momentum strategies consistently outperform the benchmarks, including a buy-and-hold strategy and the pure

---

<sup>27</sup>In the sample period 1996-2015, this strategy generates 12 transactions at all trading frequencies ranging from 1, 5, 10 to 20 days.

momentum strategy based on S&P 500 index alone. The empirical results indicate that the bubble estimates derived from option prices contain valuable forward-looking information.

The current method has some limitations. Our proposed methods are static in the sense that the state-price distribution (and hence the bubbles) are inferred from the cross-section of option data on the same day. According to the local-martingale theory, the bubble process satisfies some dynamic properties, e.g., it is a supermartingale from the birth to the death of a bubble cycle. An empirical investigation of the time series dynamics of asset price bubbles would shed light on the bubble formation process. In our empirical analysis, option prices were taken as the simple average of the bid and offer prices. The estimation method could be improved by incorporating additional data such as the bid-offer spread, so that the level of noise associated with the option price is explicitly taken into account. Furthermore, other weighting schemes could be considered when estimating the bubbles using put and call data (e.g., using a volume-weighted average in expression (16)). The relaxation of these limitations is left for future research.

#### **Conflict-of-interest disclosure statement**

Both authors have nothing to disclose.

## **References**

- [1] Aït-Sahalia, Y., and A.W. Lo (1998): Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *Journal of Finance*, 53(2), 499–547.
- [2] Aït-Sahalia, Y., and J. Duarte (2003): Nonparametric Option Pricing under Shape Restrictions, *Journal of Econometrics*, 116, 9–47.
- [3] Battalio, R., and P. Schultz (2011): Regulatory Uncertainty and Market Liquidity: The 2008 Short Sale Ban’s Impact on Equity Option Markets, *Journal of Finance*, 66(6), 2013–2053.
- [4] Black, F., and M. Scholes (1973): The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81(3), 637–654.
- [5] Breeden, D., and R. Litzenberger (1978): Prices of State Contingent Claims Implicit in Option Prices, *Journal of Business*, 51, 621–657.

- [6] Conlon, J.R. (2004): Simple Finite Horizon Bubbles Robust to Higher Order Knowledge, *Econometrica*, 72(3), 927–936.
- [7] Cox, A.M.G., and D.G. Hobson (2005): Local Martingales, Bubbles and Option Prices, *Finance and Stochastics*, 9(4), 477–492.
- [8] Delbaen, F., and W. Schachermayer (1994): A General Version of the Fundamental Theorem of Asset Pricing, *Mathematische Annalen*, 300(3), 463–520.
- [9] Delbaen, F., and W. Schachermayer (1998): The Fundamental Theorem of Asset Pricing for Unbounded Stochastic Processes, *Mathematische Annalen*, 312(2), 215–250.
- [10] Fan, J., and I. Gijbels (1996): *Local Polynomial Modelling and its Applications*, Chapman & Hall.
- [11] Fan, J., and Q. Yao (2005): *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer.
- [12] Fan, J., and L. Mancini (2009): Option Pricing With Model-Guided Nonparametric Methods, *Journal of the American Statistical Association*, 104(488), 1351–1372
- [13] Figlewski, S. (2010): Estimating the Implied Risk-Neutral Density for the US Market Portfolio. In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, Chapter 15.
- [14] Fusari, N., R.A. Jarrow, and S. Lamichhane (2020): Testing for Asset Price Bubbles using Options Data, *Working Paper*.
- [15] Greene, W. (2003): *Econometric Analysis*, Fifth Ed., Prentice Hall.
- [16] Heston, S.L., M. Loewenstein, and G.A. Willard (2006): Options and Bubbles, *Review of Financial Studies*, 20(2), 359–390.
- [17] Jarrow, R.A. (2015): Asset Price Bubbles. *Annual Review of Financial Economics*, 07, 201–218.
- [18] Jarrow, R.A., Y. Kchia, and P. Protter (2011a): How to Detect an Asset Bubble, *SIAM Journal of Financial Mathematics*, 2, 839–865.

- [19] Jarrow, R.A., Y. Kchia, and P. Protter (2011b): Is There a Bubble in LinkedIn’s Stock Price? *Journal of Portfolio Management*, 38(1), 125–130.
- [20] Jarrow, R.A., and M. Larsson (2012): The Meaning of Market Efficiency, *Mathematical Finance*, 22(1), 1–30.
- [21] Jarrow, R.A., and P. Protter (2010): The Martingale Theory of Bubbles: Implications for the Valuation of Derivatives and Detecting Bubbles. *Johnson School Research Paper Series* No. 25-2010.
- [22] Jarrow, R.A., P. Protter, and K. Shimbo (2006): Asset Price Bubbles in a Complete Market, in *Advances of Mathematical Finance, In Honor of Dilip B. Madan*, ed. by M.C. Fu, D. Madan, pp. 97–121.
- [23] Jarrow, R.A., P. Protter, and K. Shimbo (2010): Asset Price Bubbles in Incomplete Markets, *Mathematical Finance*, 20(2), 145–185.
- [24] Kitsul, Y, and J.H. Wright (2013): The Economics of Options-Implied Inflation Probability Density Functions, *Journal of Financial Economics*, 110(3), 696–711.
- [25] Loewenstein, M., and G. Willard (2000a): Rational Equilibrium Asset-Pricing Bubbles in Continuous Trading Models. *Journal of Economic Theory*, 91(1), 17–58.
- [26] Loewenstein, M., and G. Willard (2000b): Local Martingales, Arbitrage and Viability: Free Snacks and Cheap Thrills. *Economic Theory*, 16(1), 135–161.
- [27] Lu, J., and Z. Qu (2018): Sieve Estimation of Options Implied State Price Density. *Working Paper*.
- [28] Masry, E., and J. Fan (1997): Local Polynomial Estimation of Regression Functions for Mixing Processes. *Scandinavian Journal of Statistics*, 24, 165–179.
- [29] Merton, R.C. (1973): Theory of Rational Option Pricing, *Bell Journal of Economics*, 4(1), 141–183.
- [30] Milunovich, G., S. Shi, and D. Tan (2019): Bubble Detection and Sector Trading in Real Time, *Quantitative Finance*, 19(2), 247–263.



- [31] Nishiotis, G.P., and L.S. Rompolis (2019): Put-call Parity Violations and Return Predictability: Evidence from the 2008 Short Sale Ban, *Journal of Banking and Finance*, 106, 276–297.
- [32] Phillips, P.C.B., S. Shi, and J. Yu (2015a): Testing for Multiple Bubbles: Historical Episodes of Irrational Exuberance and Collapse in the S&P 500, *International Economic Review*, 56(4), 1043–1077.
- [33] Phillips, P.C.B., S. Shi, and J. Yu (2015b): Testing for Multiple Bubbles: Limit Theory of Real-time Detectors, *International Economic Review*, 56(4), 1079–1133.
- [34] Protter, P. (2005): *Stochastic Integration and Differential Equations*, Second Ed., Springer-Verlag, New York.
- [35] Protter, P. (2013): A Mathematical Theory of Financial Bubbles, in *Paris-Princeton Lectures on Mathematical Finance 2013, Lecture Notes in Mathematics 2081*, ed. by V. Henderson, R. Sircar, pp.1–108.
- [36] Shimko, D. (1993): Bounds of Probability, *Risk*, 6(4), 33–37.
- [37] Shiryaev, A.N., M.V. Zhitlukhin, and W.T. Ziemba (2014): When to Sell Apple and the NASDAQ? Trading Bubbles with a Stochastic Disorder Model, *Journal of Portfolio Management*, 40(2), 54–63.
- [38] Shiryaev, A.N., M.V. Zhitlukhin, and W.T. Ziemba (2015): Land and stock bubbles, crashes and exit strategies in Japan circa 1990 and in 2013, *Quantitative Finance*, 15(9), 1449–1469.
- [39] Song, Z., and D. Xiu (2016): A tale of two option markets: Pricing kernels and volatility risk, *Journal of Econometrics*, 190(1), 176–196.
- [40] White. H. (2000): A Reality Check for Data Snooping, *Econometrica*, 68(5), 1097–1126.
- [41] Yatchew, A., and W. Härdle (2006): Nonparametric State Price Density Estimation Using Constrained Least Squares and the Bootstrap, *Journal of Econometrics*, 133(2), 579–599.

## Figures and Tables

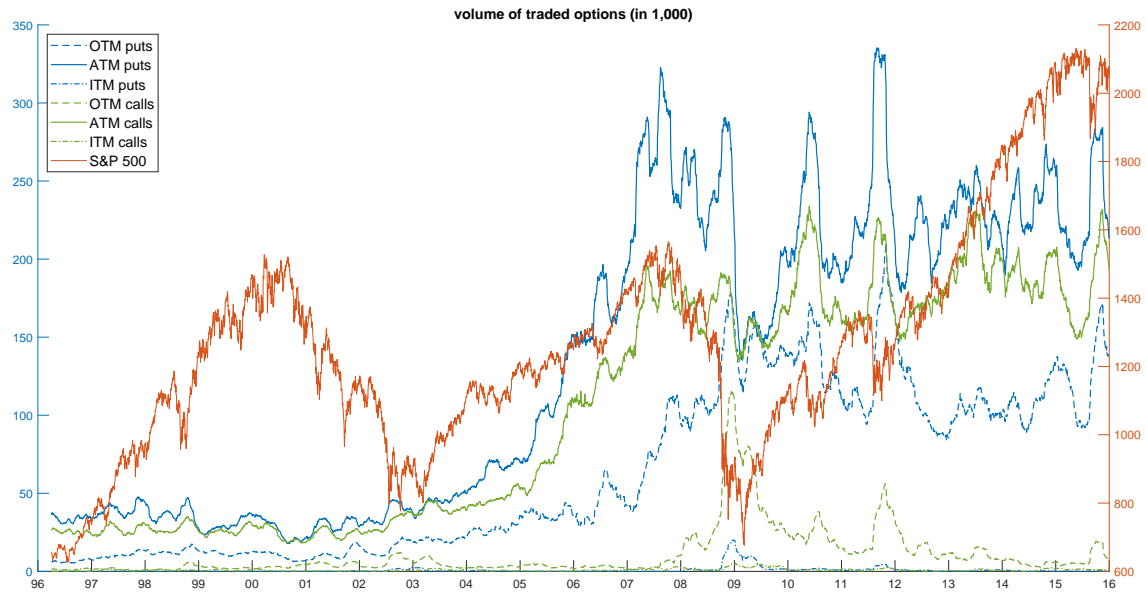


Figure 1: Trading volume (in 1,000) of calls and puts.

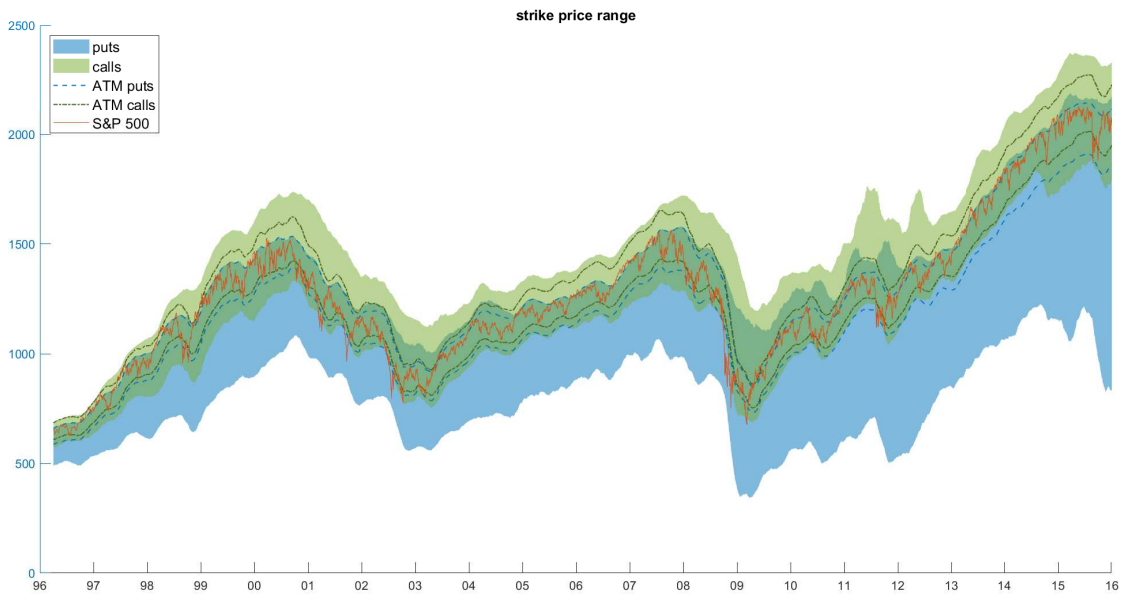


Figure 2: Range of strike prices for all calls and puts, and for at-the-money options with absolute log-moneyness less than 0.1 (boundaries in dashed and dash-dotted lines).

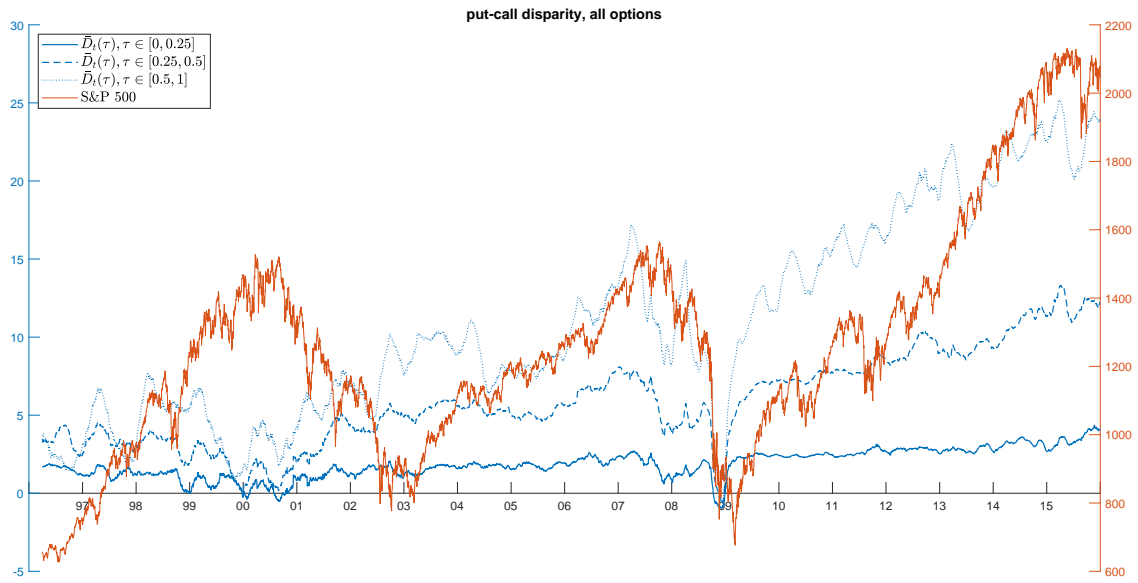


Figure 3: Put-call disparity averaged over strike prices.

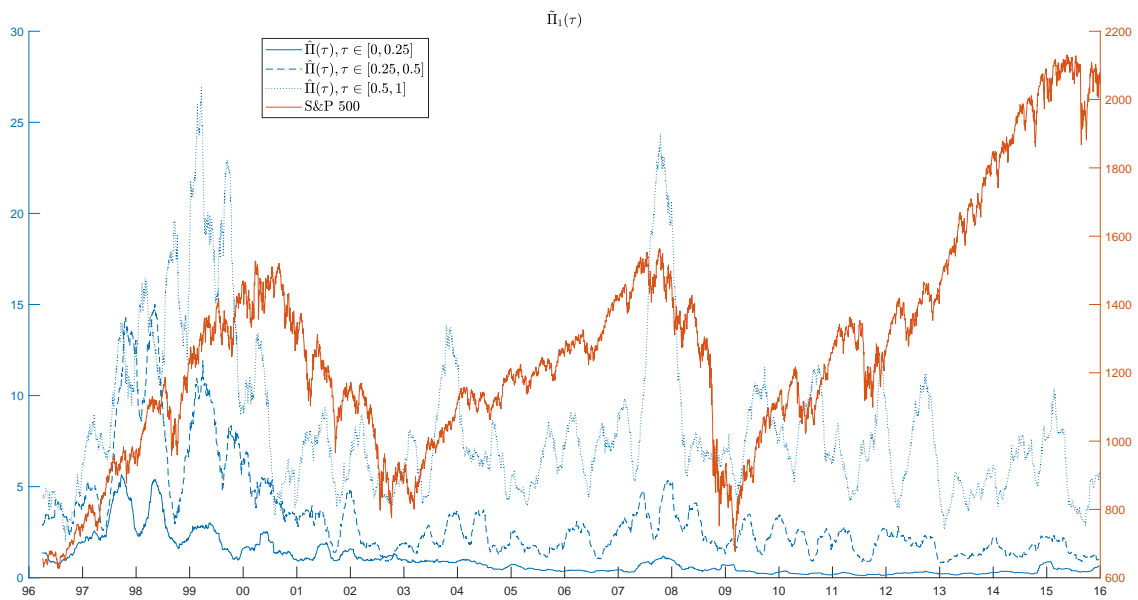


Figure 4: Naive upper bound of S&P 500 bubble obtained from the most OTM call.

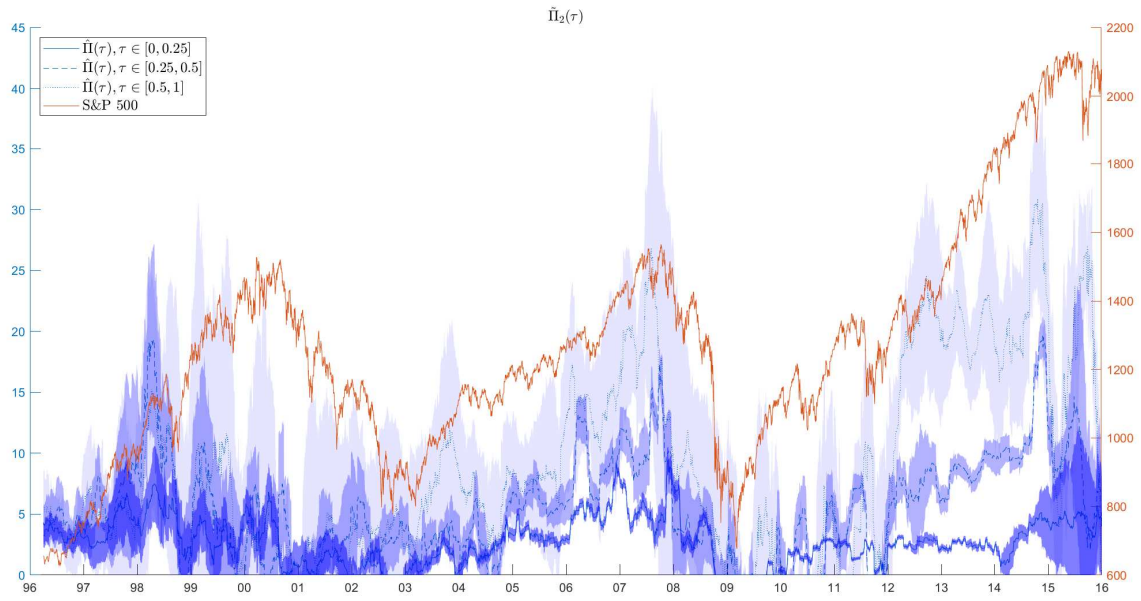


Figure 5: Simple bounds of S&P 500 bubble obtained from calls and puts of extreme monyness.

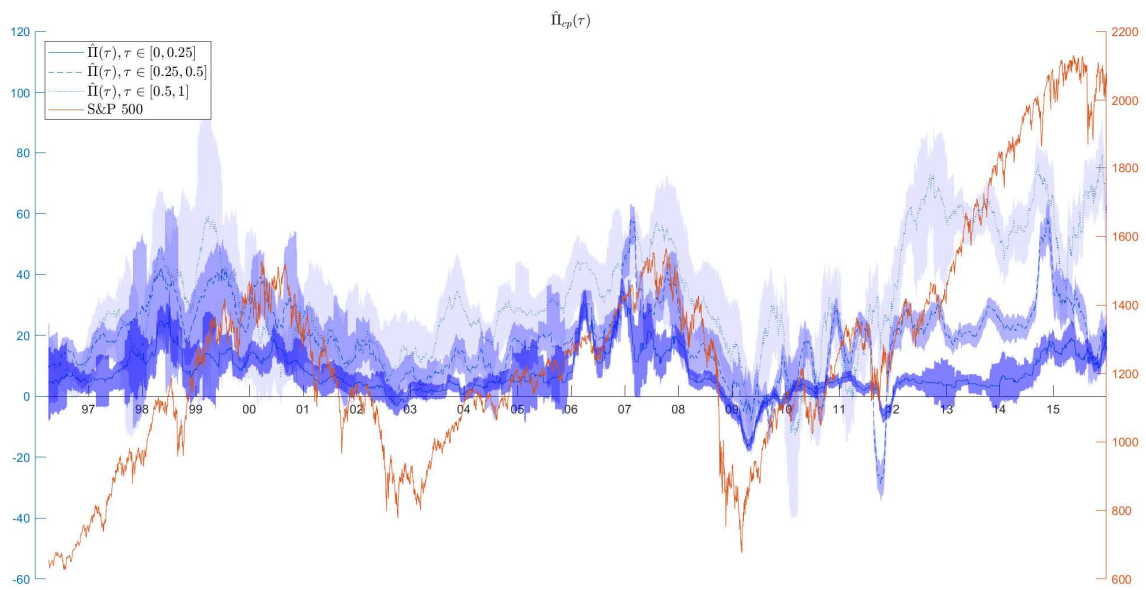


Figure 6: S&P 500 bubble estimate  $\hat{\Pi}_{cp}(\tau)$ .

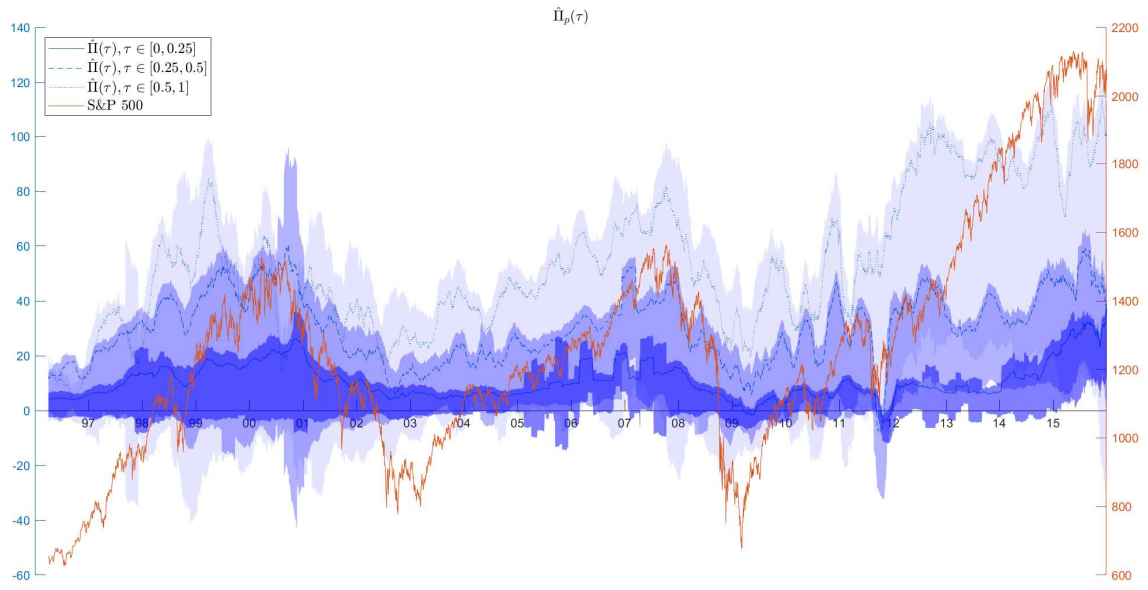


Figure 7: S&P 500 bubble estimate  $\hat{\Pi}_p(\tau)$ .

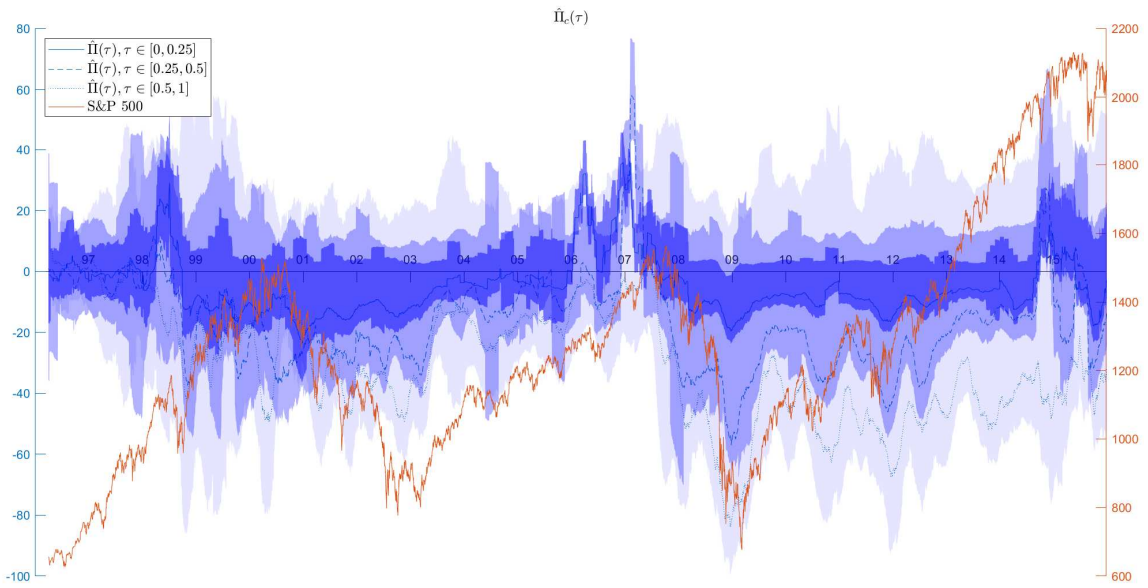


Figure 8: S&P 500 bubble estimate  $\hat{\Pi}_c(\tau)$ .

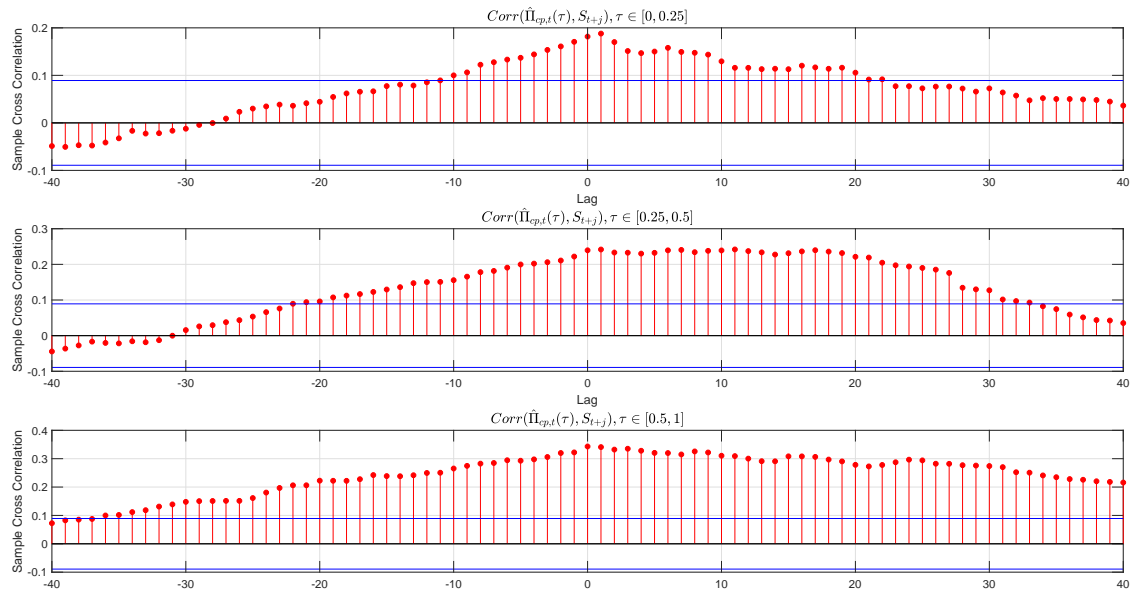


Figure 9: Cross-correlations of S&P 500 and its bubble estimate  $\hat{\Pi}_{cp}(\tau)$ .

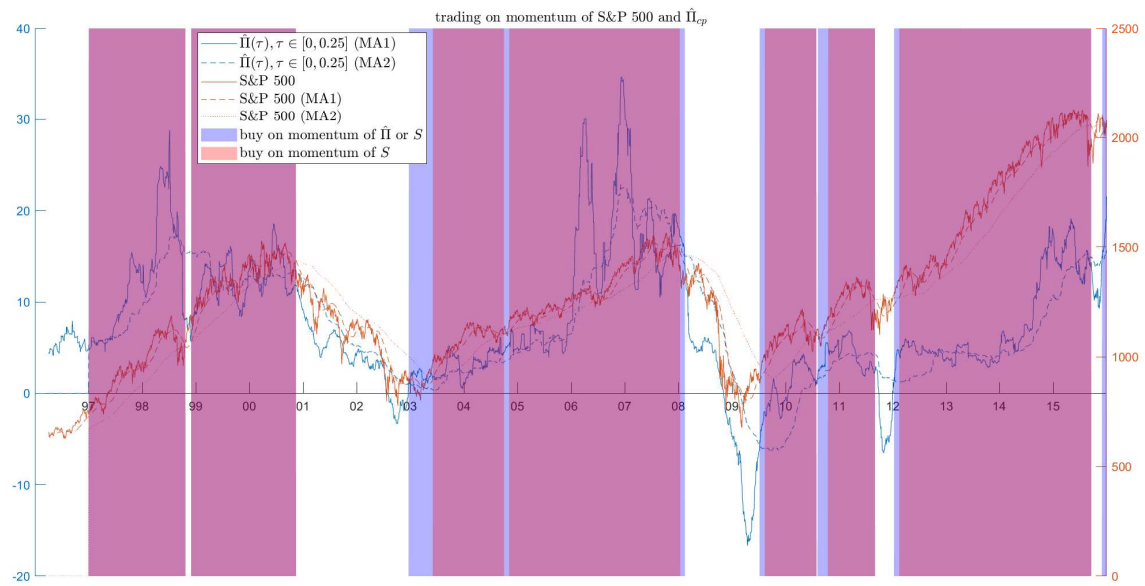


Figure 10: Momentum strategies based on S&P 500 and its bubble estimate  $\hat{\Pi}_{cp}(\tau)$  with  $\tau \in [0, 0.25]$ .

Table 1: Average daily trading volume (in 1,000) of S&P 500 index options

	log-moneyness	maturity group		
		$\tau \in [0, 0.25]$	$\tau \in [0.25, 0.5]$	$\tau \in [0.5, 1]$
OTM puts	$m \in (-\infty, -0.1)$	48.9	8.9	5.8
ATM puts	$m \in [-0.1, 0.1]$	116.3	13.0	5.4
ITM puts	$m \in (0.1, \infty)$	0.5	0.2	0.2
OTM calls	$m \in (0.1, \infty)$	6.3	2.3	2.2
ATM calls	$m \in [-0.1, 0.1]$	90.8	9.4	4.2
ITM calls	$m \in (-\infty, -0.1)$	0.6	0.2	0.1

Note: The sample horizon is 1996-2015. Log-moneyness is defined as  $m = \log(ke^{-r\tau}/S)$ .

Table 2: Performance of trading strategies on S&P 500 over different sample horizons

	1996-2015			1996-2000			2001-2005			2006-2010			2011-2015		
	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe
momentum on S&P 500 or $\Pi$															
for $\tau \in [0, 0.25]$ :															
$\hat{\Pi}_{cp}(\tau)$	9.1	8.7	0.50	13.6	13.3	0.44	6.7	6.6	0.48	9.4	9.0	0.55	6.6	6.2	0.44
$\hat{\Pi}_p(\tau)$	8.4	7.9	0.43	12.7	12.4	0.39	6.3	5.5	0.36	7.1	6.7	0.39	7.6	7.2	0.50
$\hat{\Pi}_c(\tau)$	6.5	6.1	0.27	13.6	13.3	0.44	-1.5	-1.8	-0.28	9.9	9.5	0.60	5.5	5.2	0.37
$\tilde{\Pi}_2(\tau)$	6.3	5.8	0.26	13.6	13.3	0.44	-0.9	-1.8	-0.29	8.9	8.5	0.50	6.7	6.3	0.44
$\tilde{\Pi}_1(\tau)$	5.8	5.2	0.21	13.6	13.3	0.44	4.1	3.0	0.09	2.4	1.8	-0.02	4.6	4.0	0.28
for $\tau \in [0.25, 0.5]$ :															
$\hat{\Pi}_{cp}(\tau)$	8.2	7.6	0.42	13.6	13.3	0.44	4.5	3.8	0.20	5.9	5.3	0.29	7.1	6.8	0.47
$\hat{\Pi}_p(\tau)$	8.2	7.7	0.43	13.6	13.3	0.44	5.0	4.5	0.28	6.8	6.2	0.37	7.9	7.2	0.50
$\hat{\Pi}_c(\tau)$	6.2	5.4	0.23	13.5	13.0	0.43	1.3	0.1	-0.16	5.1	4.3	0.15	5.7	5.4	0.38
$\tilde{\Pi}_2(\tau)$	6.1	5.3	0.23	13.6	13.3	0.44	0.4	-0.5	-0.22	2.2	1.0	-0.09	7.1	6.8	0.47
$\tilde{\Pi}_1(\tau)$	8.0	7.2	0.37	13.6	13.3	0.44	5.1	3.8	0.18	5.3	4.3	0.16	7.8	7.3	0.49
for $\tau \in [0.5, 1]$ :															
$\hat{\Pi}_{cp}(\tau)$	8.5	7.8	0.41	12.7	12.4	0.39	3.4	1.8	-0.02	8.7	8.3	0.55	9.9	9.9	0.63
$\hat{\Pi}_p(\tau)$	8.3	7.5	0.39	16.2	15.6	0.55	1.8	0.6	-0.13	8.6	8.2	0.51	7.9	7.5	0.48
$\hat{\Pi}_c(\tau)$	6.5	5.7	0.25	12.7	12.4	0.39	-0.2	-1.8	-0.32	7.5	7.1	0.41	5.8	5.2	0.36
$\tilde{\Pi}_2(\tau)$	7.5	6.9	0.32	13.6	13.3	0.44	1.6	0.3	-0.12	8.7	8.3	0.55	8.0	7.3	0.61
$\tilde{\Pi}_1(\tau)$	7.2	6.6	0.31	13.6	13.3	0.44	2.0	0.9	-0.10	5.7	5.3	0.24	8.4	8.0	0.52
momentum on S&P 500	7.3	6.9	0.39	13.6	13.3	0.44	4.1	3.8	0.24	5.7	5.3	0.29	5.5	5.2	0.37
buy & hold	5.1	5.1	0.16	11.7	11.7	0.34	-0.5	-0.5	-0.19	-0.1	-0.1	-0.12	7.9	7.9	0.50

Note: All reported returns are annualized and in percentage. Sharpe ratios are annualized. Net returns are computed after accounting for transaction cost (0.05% per round-trip transaction). Trading frequency is once a week. The second-to-last row corresponds to momentum strategy based on S&P 500 only. The last row corresponds to the buy-and-hold strategy. The rest of the rows correspond to momentum strategies based on S&P 500 or its various bubble estimates with time-to-liquidation  $\tau$ . All momentum strategies yield a hold signal when the short-span (3 months) moving average exceeds the long-span (1 year) moving average of the associated variable. See the main text for definitions of the bubble estimates.



Table 3: Performance of trading strategies on S&P 500 for different trading frequencies

	$\Delta t$ : 1 day			$\Delta t$ : 5 days			$\Delta t$ : 10 days			$\Delta t$ : 20 days		
	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe
momentum on S&P 500 or $\Pi$												
for $\tau \in [0, 0.25]$ :												
$\hat{\Pi}_{cp}(\tau)$	8.6	8.2	0.43	9.1	8.7	0.50	8.8	8.5	0.52	7.8	7.5	0.48
$\hat{\Pi}_p(\tau)$	8.2	7.4	0.37	8.4	7.9	0.43	7.8	7.3	0.44	7.9	7.5	0.47
$\hat{\Pi}_c(\tau)$	6.3	5.7	0.24	6.5	6.1	0.27	6.5	6.2	0.31	6.4	6.0	0.31
$\tilde{\Pi}_2(\tau)$	6.7	6.0	0.26	6.3	5.8	0.26	7.1	6.7	0.35	6.1	5.7	0.29
$\tilde{\Pi}_1(\tau)$	6.0	5.2	0.20	5.8	5.2	0.21	6.1	5.5	0.26	6.0	5.4	0.25
for $\tau \in [0.25, 0.5]$ :												
$\hat{\Pi}_{cp}(\tau)$	7.9	7.3	0.37	8.2	7.6	0.42	7.7	7.2	0.43	6.8	6.4	0.38
$\hat{\Pi}_p(\tau)$	8.2	7.6	0.39	8.2	7.7	0.43	7.7	7.2	0.43	7.4	7.0	0.43
$\hat{\Pi}_c(\tau)$	6.5	5.4	0.20	6.2	5.4	0.23	5.3	4.7	0.18	5.8	5.2	0.24
$\tilde{\Pi}_2(\tau)$	5.7	4.7	0.18	6.1	5.3	0.23	6.0	5.4	0.26	5.9	5.4	0.28
$\tilde{\Pi}_1(\tau)$	7.5	6.4	0.29	8.0	7.2	0.37	7.3	6.6	0.35	7.7	7.2	0.42
for $\tau \in [0.5, 1]$ :												
$\hat{\Pi}_{cp}(\tau)$	9.1	8.0	0.38	8.5	7.8	0.41	7.9	7.4	0.42	7.2	6.7	0.37
$\hat{\Pi}_p(\tau)$	8.0	6.8	0.31	8.3	7.5	0.39	8.2	7.5	0.41	7.7	7.2	0.41
$\hat{\Pi}_c(\tau)$	6.9	5.7	0.24	6.5	5.7	0.25	6.4	5.8	0.29	5.3	4.8	0.21
$\tilde{\Pi}_2(\tau)$	8.0	6.9	0.30	7.5	6.9	0.32	6.9	6.3	0.31	6.6	6.1	0.31
$\tilde{\Pi}_1(\tau)$	6.9	6.1	0.26	7.2	6.6	0.31	6.8	6.3	0.31	6.5	6.0	0.31
momentum on S&P 500	7.0	6.7	0.34	7.3	6.9	0.39	6.5	6.1	0.35	6.2	5.8	0.34
buy & hold	5.1	5.1	0.14	5.1	5.1	0.16	5.1	5.1	0.17	5.1	5.1	0.17

Note: All reported returns are annualized and in percentage. Sharpe ratios are annualized. Net returns are computed after accounting for transaction cost (0.05% per round-trip transaction). Trading frequency is every  $\Delta t$  days (vary from 1 day to 20 days). The second-to-last row corresponds to momentum strategy based on S&P 500 only. The last row corresponds to the buy-and-hold strategy. The rest of the rows correspond to momentum strategies based on S&P 500 or its various bubble estimates with time-to-liquidation  $\tau$ . All momentum strategies yield a hold signal when the short-span (3 months) moving average exceeds the long-span (1 year) moving average of the associated variable. See the main text for definitions of the bubble estimates. The sample horizon is 1996-2015.

Table 4: Performance of trading strategies on S&P 500 for different spans of moving averages

	MA1: 3 mo MA2: 1 yr			MA1: 1 mo MA2: 1 yr			MA1: 6 mo MA2: 1 yr			MA1: 3 mo MA2: 6 mo			MA1: 3 mo MA2: 2 yr		
	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe
momentum on S&P 500 or $\Pi$															
for $\tau \in [0, 0.25]$ :															
$\hat{\Pi}_{cp}(\tau)$	9.1	8.7	0.50	8.6	7.6	0.42	8.7	8.5	0.48	5.2	4.0	0.13	7.3	7.0	0.38
$\hat{\Pi}_p(\tau)$	8.4	7.9	0.43	7.8	6.9	0.37	7.5	7.2	0.38	6.0	5.2	0.22	6.4	6.2	0.33
$\hat{\Pi}_c(\tau)$	6.5	6.1	0.27	5.5	4.5	0.17	7.1	6.8	0.31	6.8	5.7	0.25	6.1	5.7	0.25
$\tilde{\Pi}_2(\tau)$	6.3	5.8	0.26	6.2	5.0	0.21	7.5	6.9	0.32	3.3	2.1	-0.01	5.6	5.2	0.24
$\tilde{\Pi}_1(\tau)$	5.8	5.2	0.21	6.5	5.3	0.23	7.0	6.5	0.30	3.0	1.8	-0.03	4.8	4.4	0.17
for $\tau \in [0.25, 0.5]$ :															
$\hat{\Pi}_{cp}(\tau)$	8.2	7.6	0.42	7.2	6.1	0.31	7.8	7.4	0.40	6.1	5.2	0.23	6.6	6.4	0.34
$\hat{\Pi}_p(\tau)$	8.2	7.7	0.43	9.0	7.8	0.44	8.6	8.4	0.47	6.1	5.1	0.22	6.7	6.6	0.35
$\hat{\Pi}_c(\tau)$	6.2	5.4	0.23	5.6	4.4	0.16	5.3	4.6	0.16	6.5	5.4	0.22	6.0	5.5	0.24
$\tilde{\Pi}_2(\tau)$	6.1	5.3	0.23	7.0	5.5	0.24	6.2	5.7	0.25	4.5	3.5	0.08	5.2	4.7	0.19
$\tilde{\Pi}_1(\tau)$	8.0	7.2	0.37	8.0	6.7	0.33	9.4	9.0	0.50	7.4	6.0	0.26	6.0	5.5	0.26
for $\tau \in [0.5, 1]$ :															
$\hat{\Pi}_{cp}(\tau)$	8.5	7.8	0.41	6.3	5.3	0.22	7.3	7.1	0.33	6.3	5.3	0.22	6.1	5.4	0.26
$\hat{\Pi}_p(\tau)$	8.3	7.5	0.39	7.5	6.1	0.28	7.9	7.6	0.39	6.6	5.3	0.21	7.3	7.0	0.37
$\hat{\Pi}_c(\tau)$	6.5	5.7	0.25	5.6	4.6	0.17	7.2	6.8	0.32	5.0	3.9	0.11	7.0	6.6	0.34
$\tilde{\Pi}_2(\tau)$	7.5	6.9	0.32	4.9	3.7	0.11	6.2	5.8	0.24	5.3	4.1	0.13	4.9	4.3	0.15
$\tilde{\Pi}_1(\tau)$	7.2	6.6	0.31	6.0	4.7	0.18	9.4	8.9	0.46	5.7	4.6	0.15	6.6	6.3	0.31
momentum on S&P 500	7.3	6.9	0.39	6.8	6.2	0.33	7.3	7.1	0.38	5.9	5.1	0.23	6.3	6.1	0.32
buy & hold	5.1	5.1	0.16	5.1	5.1	0.16	5.1	5.1	0.16	5.6	5.6	0.18	3.7	3.7	0.07

Note: All reported returns are annualized and in percentage. Sharpe ratios are annualized. Net returns are computed after accounting for transaction cost (0.05% per round-trip transaction). Trading frequency is once a week. The second-to-last row corresponds to momentum strategy based on S&P 500 only. The last row corresponds to the buy-and-hold strategy. The rest of the rows correspond to momentum strategies based on S&P 500 or its various bubble estimates with time-to-liquidation  $\tau$ . All momentum strategies yield a hold signal when the short-span moving average (MA1) exceeds the long-span moving average (MA2) of the associated variable. See the main text for definitions of the bubble estimates. The sample horizon is 1996-2015.

Table 5: Reality Check tests

Average block size	H0: buy-and-hold strategy H1: Joint momentum strategy based on S&P 500 and $\hat{\Pi}_{cp}(\tau)$			H0: buy-and-hold strategy H1: Pure momentum strategy based on S&P 500
	maturity group			
	$\tau \in [0, 0.25]$	$\tau \in [0.25, 0.5]$	$\tau \in [0.5, 1]$	
3 months	0.076	0.074	0.060	0.105
6 months	0.018	0.026	0.015	0.024
9 months	0.019	0.014	0.012	0.015
12 months	0.015	0.014	0.007	0.012

Note: The sample horizon is 1996-2015. The p-value of Bootstrap Reality Check test of White (2000) are reported for each given average block size used for the stationary bootstrap (first column). The benchmark is the buy-and-hold strategy. The tests for the superiority of joint momentum strategy (columns 2-4) take into account different combinations of moving average window lengths ( $w_1=21,42,63,84,105,126$  days,  $w_2=126,252,378,504$  days, where  $w_1 < w_2$ ) and trading frequencies (1,5,10,20 days). The tests for the superiority of pure momentum strategy (last column) take into account different trading frequencies (1,5,10,20 days). Number of bootstrap samples: 1,000.

Table 6: Performance of different variants of trading strategies on S&P 500

	$\Pi$ or S&P 500			$\Pi$ only			$\Pi$ and S&P 500		
	gross	net	Sharpe	gross	net	Sharpe	gross	net	Sharpe
choice of $\Pi$									
for $\tau \in [0, 0.25]$ :									
$\hat{\Pi}_{cp}(\tau)$	9.1	8.7	0.50	5.6	4.2	0.19	3.8	2.4	0.02
$\hat{\Pi}_p(\tau)$	8.4	7.9	0.43	4.7	3.7	0.13	3.6	2.7	0.05
$\hat{\Pi}_c(\tau)$	6.5	6.1	0.27	3.1	2.1	-0.01	3.9	2.9	0.09
$\tilde{\Pi}_2(\tau)$	6.3	5.8	0.26	2.7	0.7	-0.13	3.7	1.8	-0.04
$\tilde{\Pi}_1(\tau)$	5.8	5.2	0.21	2.3	0.8	-0.13	3.8	2.6	0.05
for $\tau \in [0.25, 0.5]$ :									
$\hat{\Pi}_{cp}(\tau)$	8.2	7.6	0.42	4.6	3.3	0.10	3.7	2.6	0.04
$\hat{\Pi}_p(\tau)$	8.2	7.7	0.43	6.2	4.9	0.25	5.3	4.2	0.18
$\hat{\Pi}_c(\tau)$	6.2	5.4	0.23	2.0	0.5	-0.15	3.1	2.0	-0.02
$\tilde{\Pi}_2(\tau)$	6.1	5.3	0.23	2.5	0.7	-0.13	3.7	2.3	0.02
$\tilde{\Pi}_1(\tau)$	8.0	7.2	0.37	3.0	1.5	-0.07	2.3	1.2	-0.12
for $\tau \in [0.5, 1]$ :									
$\hat{\Pi}_{cp}(\tau)$	8.5	7.8	0.41	4.8	3.2	0.09	3.7	2.3	0.01
$\hat{\Pi}_p(\tau)$	8.3	7.5	0.39	5.2	3.6	0.12	4.3	3.0	0.08
$\hat{\Pi}_c(\tau)$	6.5	5.7	0.25	0.8	-0.8	-0.27	1.6	0.4	-0.21
$\tilde{\Pi}_2(\tau)$	7.5	6.9	0.32	4.2	2.6	0.03	4.0	2.6	0.05
$\tilde{\Pi}_1(\tau)$	7.2	6.6	0.31	3.8	2.4	0.02	3.9	2.8	0.07
momentum on S&P 500	7.3	6.9	0.39	7.3	6.9	0.39	7.3	6.9	0.39
buy & hold	5.1	5.1	0.16	5.1	5.1	0.16	5.1	5.1	0.16

Note: All reported returns are annualized and in percentage. Sharpe ratios are annualized. Net returns are computed after accounting for transaction cost (0.05% per round-trip transaction). Trading frequency is once a week. The second-to-last row corresponds to momentum strategy based on S&P 500 only. The last row corresponds to the buy-and-hold strategy. The rest of the rows correspond to momentum strategies based on different signalling variables: S&P 500 and/or its bubble estimate with time-to-liquidation  $\tau$ , or the bubble estimate alone. All momentum strategies yield a hold signal when the short-span (3 months) moving average exceeds the long-span (1 year) moving average of the signalling variable. See the main text for definitions of the bubble estimates. The sample horizon is 1996-2015.

Table 7: Performance of trading strategies on S&P 500 under different transaction costs

	0%	0.5%	1%	2%	3%
momentum on S&P 500 or $\Pi$					
for $\tau \in [0, 0.25]$ :					
$\hat{\Pi}_{cp}(\tau)$	9.1	8.7	8.4	7.8	7.2
$\hat{\Pi}_p(\tau)$	8.4	7.9	7.4	6.5	5.6
$\hat{\Pi}_c(\tau)$	6.5	6.1	5.7	4.9	4.1
$\tilde{\Pi}_2(\tau)$	6.3	5.8	5.4	4.4	3.5
$\tilde{\Pi}_1(\tau)$	5.8	5.2	4.6	3.3	2.0
for $\tau \in [0.25, 0.5]$ :					
$\hat{\Pi}_{cp}(\tau)$	8.2	7.6	7.1	5.9	4.8
$\hat{\Pi}_p(\tau)$	8.2	7.7	7.1	6.0	4.9
$\hat{\Pi}_c(\tau)$	6.2	5.4	4.7	3.1	1.5
$\tilde{\Pi}_2(\tau)$	6.1	5.3	4.4	2.8	1.2
$\tilde{\Pi}_1(\tau)$	8.0	7.2	6.5	5.0	3.4
for $\tau \in [0.5, 1]$ :					
$\hat{\Pi}_{cp}(\tau)$	8.5	7.8	7.2	5.9	4.7
$\hat{\Pi}_p(\tau)$	8.3	7.5	6.8	5.3	3.8
$\hat{\Pi}_c(\tau)$	6.5	5.7	5.0	3.5	2.0
$\tilde{\Pi}_2(\tau)$	7.5	6.9	6.3	5.1	3.8
$\tilde{\Pi}_1(\tau)$	7.2	6.6	6.0	4.9	3.7
momentum on S&P 500	7.3	6.9	6.6	5.9	5.2
buy & hold	5.1	5.1	5.1	5.1	5.1

Note: All reported returns are annualized and in percentage. Sharpe ratios are annualized. Net returns are computed after accounting for transaction cost (vary from 0% to 3% per round-trip transaction). Trading frequency is once a week. The second-to-last row corresponds to momentum strategy based on S&P 500 only. The last row corresponds to the buy-and-hold strategy. The rest of the rows correspond to momentum strategies based on S&P 500 or its various bubble estimates with time-to-liquidation  $\tau$ . All momentum strategies yield a hold signal when the short-span moving average (MA1) exceeds the long-span moving average (MA2) of the associated variable. See the main text for definitions of the bubble estimates. The sample horizon is 1996-2015.

# Appendix

## A1 Details of the Bubble Estimation Procedure

We now discuss the problem of recovering the state-price distribution from observed option data. The estimation procedure, to be described below, delivers an asset price bubble estimate that respects the data constraints (bounded strike price range, heterogeneous data precision) and is consistent with Assumptions NFLVR and ND. It involves an application of constrained least squares and local polynomial fitting.

**Step 1: Constrained Least Squares** Suppose each cross-sectional sample consists of put or call prices over different strikes with fixed time  $t$  and time-to-maturity  $\tau$  (we suppress the subscripts  $t$  and  $\tau$  for notational simplicity). As our first step of the estimation procedure, we mitigate the data issue by conducting a constrained least squares estimation on each cross-sectional sample of option prices over strikes. More precisely, we solve a minimization problem subject to monotonicity and convexity requirements. For a given cross-sectional sample of put prices  $\{k_i, P_i\}_{i=1}^n$ , we solve:

$$\begin{aligned} & \min_{\{m_i\}_{i=1}^n \in \mathbb{R}^n} \sum_{i=1}^n (m_i - P_i)^2 \\ \text{subject to } & 0 \leq \frac{m_{i+1} - m_i}{k_{i+1} - k_i} \leq e^{-r\tau} \text{ for } i = 1, \dots, n-1 \text{ (monotonicity)} \\ & \text{and } \frac{m_{i+1} - m_i}{k_{i+1} - k_i} \leq \frac{m_{i+2} - m_{i+1}}{k_{i+2} - k_{i+1}} \text{ for } i = 1, \dots, n-2 \text{ (convexity)}. \end{aligned}$$

For call price data  $\{k_i, C_i\}_{i=1}^n$ , the optimization problem is similar, except that the input data are call prices and the monotonicity constraints are changed to  $-e^{-r\tau} \leq \frac{m_{i+1} - m_i}{k_{i+1} - k_i} \leq 0$  for  $i = 1, \dots, n-1$ . The solution  $\{m_i\}_{i=1}^n$  obeys the monotonicity and convexity restrictions.

**Step 2: Local Polynomial Fitting** As our second step, we estimate the cdf  $Q^*(s)$  of the state-price density using the adjusted option data obtained from step 1. For fixed  $t$  and  $\tau$ , we let  $m_p(s)$  and  $m_c(s)$  denote the put and call option pricing functions that respect monotonicity and convexity restrictions, and  $m_p^{(\nu)}(s)$  and  $m_c^{(\nu)}(s)$  their  $\nu^{\text{th}}$  derivatives. Recall from Lemma 1 that

$Q^*(s)$  is connected to the first-order derivative of the pricing function as follows:

$$Q_p^*(s) = e^{r\tau} m_p^{(1)}(k) \quad \text{if put data are used, or}$$

$$Q_c^*(s) = e^{r\tau} m_c^{(1)}(k) + 1 \quad \text{if call data are used.}$$

This motivates the following estimators of  $Q^*(s)$ :

$$\hat{Q}_p^*(s) := e^{r\tau} \hat{m}_p^{(1)}(s),$$

$$\hat{Q}_c^*(s) := e^{r\tau} \hat{m}_c^{(1)}(s) + 1,$$

where  $\hat{m}_p^{(1)}(s)$  and  $\hat{m}_c^{(1)}(s)$  are local estimators of  $m_p^{(1)}(k)$  and  $m_c^{(1)}(k)$ , respectively.

To obtain  $\hat{m}^{(1)}(s)$  and its standard error (henceforth we drop the subscripts  $c$  and  $p$  as the procedure applies to both calls and puts), we apply the nonparametric local polynomial fit (Fan and Gijbels (1996), Fan and Yao (2005)). Our data consist of pairs of strike prices and adjusted option prices  $\{x_i, y_i\}_{i=1}^n = \{k_i, m_i\}_{i=1}^n$ . Next, we fit a local polynomial function of order  $p$  to the data ( $p$  to be determined below). To estimate the functional value at  $s$ , we first obtain the  $(p+1)$ -vector  $\hat{\beta} := (X'WX)^{-1}X'Wy$ . Here,  $X$  is a  $n \times (p+1)$  design matrix whose  $(i, j)^{th}$  element is  $(k_i - s)^{j-1}$ ,  $W$  is a  $(p+1) \times (p+1)$  diagonal matrix whose  $i^{th}$  element is  $K_h(k_i - s)$  ( $K_h(x) = h^{-1}K(x/h)$ , where  $K$  is the kernel and  $h$  is the bandwidth, to be discussed below), and  $y$  is a  $(p+1)$ -vector whose  $i^{th}$  element is  $m_i$  (same notations as in Fan and Yao (2005), Section 6.3.2). The local polynomial estimator for  $m^{(\nu)}(s)$  ( $\nu = 0, 1, \dots, p$ ) is given by  $\hat{m}^{(\nu)}(s) = \nu! \hat{\beta}_\nu$ , where  $\hat{\beta}_\nu$  denotes the  $(\nu+1)^{th}$  element of  $\hat{\beta}$ . Asymptotic normality result holds for the local polynomial estimator applied to dependent data that satisfy certain stationary and mixing conditions (Masry and Fan (1993); see also Fan and Gijbels (1996), Section 6.2.2). The asymptotic theory provides guidance for the choice of the polynomial order  $p$ . To estimate  $m^{(1)}(s)$  (i.e.,  $\nu = 1$ ), the best choice is to fit a local quadratic function to the data, i.e.,  $p = \nu + 1 = 2$  (Fan and Gijbels (1996), Section 3.3; Fan and Yao (2005), Section 6.3.2). This will minimize both the model complexity and the boundary bias of the estimator. By preasymptotic substitution (Fan and Yao (2005), Section 6.3.4),

the bias and variance of  $\hat{\beta}$  evaluated at  $s$  are estimated by:

$$\begin{aligned}\hat{B}(s) &= (X'WX)^{-1}X'W\hat{r}, \\ \hat{V}(s) &= \text{diag}(\hat{\Sigma}),\end{aligned}$$

where  $\hat{\Sigma} = \hat{\sigma}^2(X'WX)^{-1}(X'WWX)(X'WX)^{-1}$ ,  $\hat{\sigma}^2 = (X'WX)^{-1}X'W\hat{r}^2$ , and  $\hat{r} = y - X\hat{\beta}$  is the residual vector, all evaluated at  $s$ . It follows that the standard error of  $\hat{m}^{(\nu)}(s)$  is given by  $s.e.(\hat{m}^{(\nu)}(s)) = \nu! \sqrt{\hat{V}_\nu(s)}$ , where  $\hat{V}_\nu(s)$  denotes the  $(\nu + 1)^{th}$  diagonal element of  $\hat{\Sigma}$ . Since the bias of  $\hat{m}^{(\nu)}(s)$  for  $m^{(\nu)}(s)$  involves higher-order derivatives, which are imprecisely estimated in finite samples, we follow the usual practice by ignoring the bias term (see Fan and Yao (2005), Section 6.3.4).

Both the point estimate  $\hat{m}^{(\nu)}(s)$  and its standard error are functions of the kernel  $K$  and the bandwidth parameter  $h$ . In our application, we adopt the Gaussian kernel with optimal local bandwidth given by (Fan and Gijbels (1996), equation (3.20), with  $\nu = 1$  and  $p = 2$ ):

$$\hat{h}_{opt}(s) = 0.884 \left[ \frac{msr(s)}{n\{\hat{m}^{(3)}(s)\}^2 \hat{f}(s)} \right]^{1/9},$$

where  $msr(s)$  denotes the mean squared residuals  $msr(s) = \frac{1}{n} \sum_{i=1}^n [m_i - \hat{m}(s)]^2$ , and  $\hat{f}(s)$  is the kernel density estimate  $\hat{f}(s) = \frac{1}{nh_0} \sum_{i=1}^n K\left(\frac{k_i - s}{h_0}\right)$ . To compute the plug-in estimates  $msr(s)$  and  $\hat{f}(s)$ , we use a pilot bandwidth  $h_0$ , which is set to be three times the average width of the strike price intervals. To obtain  $\hat{m}^{(3)}(s)$  (for computing  $\hat{h}_{opt}(s)$  only), we first fit a  $(p + 3) = 5^{th}$ -order polynomial function to the pricing function  $m(s) = \sum_{j=0}^5 \gamma_j s^j$  by least squares, then differentiate it three times and evaluate at  $s$ , i.e.,  $\hat{m}^{(3)}(s) = \sum_{j=1}^3 \frac{(j+2)!}{j!} \hat{\gamma}_{j+2} s^j$ , where  $\hat{\gamma}_j$  is the least squares estimate of  $\gamma_j$ .

**Step 3: Point and Interval Estimation of Bubbles** By the above procedure, we obtain  $\hat{m}_o^{(1)}(s)$  and hence  $\hat{Q}_o^*(s)$  ( $o = c, p$ ). Since the option data allow us to estimate  $Q^*(s)$  over a bounded range  $[\ell, u]$  only, we normalize  $\hat{Q}_o^*(s)$  by (12) to obtain an estimator  $\hat{Q}_o^\dagger(s)$  of the tail-truncated distribution  $Q^\dagger(s)$ . To estimate the fundamental asset value under  $Q^\dagger(s)$ , fix a grid of deterministic values  $\{s_i\}_{i=0}^{n_{os}}$  in the strike price range  $[\ell_o, u_o]$  with  $s_0 = \ell_o$  and  $s_{n_{os}} = u_o$ , and define



the step size  $\Delta s_i := s_i - s_{i-1}$ .<sup>28</sup> At each value in the grid, we compute  $\hat{m}_o^{(1)}(s_i)$  and its standard error  $s.e.(\hat{m}_o^{(1)}(s_i))$  as above. The weighted estimators  $\hat{Q}^\dagger(s)$  and  $\hat{\mu}^\dagger$  are as given in (14)-(17). Note that  $\hat{\mu}^\dagger = \omega \hat{\mu}_c^\dagger + (1 - \omega) \hat{\mu}_p^\dagger$  ( $\omega = [0, 1]$ ).

In our empirical applications, the weight  $\omega$  is chosen to be the proportion of traded call contracts out of all  $n$  traded options in a cross section (with fixed  $t$  and  $\tau$ ), i.e.,  $\omega = \frac{n_c}{n}$  ( $n_c$  is the number of call contracts in the data).

With the help of Lemma 3, we can correct and bound the bias of  $\hat{\mu}^\dagger$  caused by tail truncation. Define  $\hat{B}_o$  ( $o = c, p$ ) to be the estimator of  $B$  (given in Lemma 3) after replacing  $Q^*(s)$  by  $\hat{Q}_o^*(s)$ . Since  $\hat{Q}_o^*(s)$  are consistent for  $Q^*(s)$  for each  $s \in [\ell_o, u_o]$ , the estimators  $\hat{B}_o$  ( $o = c, p$ ) and hence  $\hat{B} := \omega \hat{B}_c + (1 - \omega) \hat{B}_p$  ( $\omega = [0, 1]$ ) are consistent for  $B$ . We then obtain the fundamental value  $\hat{\mu} := \hat{\mu}^\dagger - \hat{B}$  with the bias term  $B$  removed.

Both  $\hat{B}$  and  $\hat{\mu}^\dagger$  involve sample variability and contribute to the variance of the bias-adjusted fundamental asset value  $\hat{\mu} := \hat{\mu}^\dagger - \hat{B}$ . To compute the variance of  $\hat{\mu}$ , the following lemma is useful.

**Lemma A1** We have

$$(\hat{B}_o - \hat{\mu}_o^\dagger) - (B - \mu^\dagger) = e^{-r\tau} \sum_{i=1}^{n_{os}} w_{oi} [\hat{Q}_o^*(s_i) - Q^*(s_i)],$$

where the weights are given by

$$\begin{aligned} w_{oi} &= e^{-r\tau} \frac{\Delta s_i}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)} \quad \text{for } i = 2, \dots, n_{os} - 1, \\ w_{o1} &= \frac{e^{-r\tau} \sum_{i=1}^{n_{os}} Q^*(s_i) \Delta s_i - e^{-r\tau} (u_o - \ell_o) [\hat{Q}_o^*(u_o) - Q^*(u_o)] - R_o}{[\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)][Q^*(u_o) - Q^*(\ell_o)]} + e^{-r\tau} \frac{\Delta s_1}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)}, \\ w_{on_{os}} &= \frac{-e^{-r\tau} \sum_{i=1}^{n_{os}} Q^*(s_i) \Delta s_i + e^{-r\tau} (u_o - \ell_o) [\hat{Q}_o^*(\ell_o) - Q^*(\ell_o)] + R_o}{[\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)][Q^*(u_o) - Q^*(\ell_o)]} + e^{-r\tau} \frac{\Delta s_{n_s}}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)}, \end{aligned}$$

with  $R_p = P(u_p) - P(\ell_p)$  and  $R_c = C(u_c) - C(\ell_c) + e^{-r\tau}(u_c - \ell_c)$ .

<sup>28</sup>In our empirical application, the step size  $\Delta s_i$  is set to be 5. The point and interval estimation results are quite insensitive to the choice of  $\Delta s_i$ .

By the above lemma, a Newey-West standard error of  $\hat{B}_o - \hat{\mu}_o^\dagger$  ( $o = c, p$ ) is given by

$$s.e.(\hat{B}_o - \hat{\mu}_o^\dagger) = \left\{ \sum_{i=1}^{n_{os}} (s.e.[\hat{w}_{oi}\hat{m}_o^{(1)}(s_i)]\Delta s_i)^2 + 2 \sum_{l=1}^{L_o} \sum_{i=1}^{n_{os}-l} \left(1 - \frac{l}{L_o+1}\right) \hat{w}_{oi}\hat{w}_{o,i+l}\tilde{m}_o^{(1)}(s_i)\tilde{m}_o^{(1)}(s_{i+l})(\Delta s_i)(\Delta s_{i+l}) \right\}^{1/2},$$

where  $\tilde{m}_o^{(1)}(s_i) := \hat{m}_o^{(1)}(s_i) - n_{os}^{-1} \sum_{j=1}^{n_{os}} \hat{m}_o^{(1)}(s_j)$  is the demeaned estimate of  $\hat{m}_o^{(1)}(s_i)$ , and

$$\begin{aligned} \hat{w}_{oi} &= e^{-r\tau} \frac{\Delta s_i}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)} \quad \text{for } i = 2, \dots, n_{os} - 1, \\ \hat{w}_{o1} &= \frac{e^{-r\tau} \sum_{i=1}^{n_{os}} \hat{Q}_o^*(s_i)\Delta s_i - R_o}{[\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)]^2} + e^{-r\tau} \frac{\Delta s_1}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)}, \\ \hat{w}_{on_{os}} &= \frac{-e^{-r\tau} \sum_{i=1}^{n_{os}} \hat{Q}_o^*(s_i)\Delta s_i + R_o}{[\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)]^2} + e^{-r\tau} \frac{\Delta s_{n_{os}}}{\hat{Q}_o^*(u_o) - \hat{Q}_o^*(\ell_o)}, \end{aligned}$$

with  $R_c = C(u_c) - C(\ell_c) + e^{-r\tau}(u_c - \ell_c)$  and  $R_p = P(u_p) - P(\ell_p)$ . The maximum lag order  $L_o$  is chosen to be the smallest integer greater than or equal to  $n_{os}^{1/4}$  (see Greene (2003)).

Assuming independence of the pricing error of calls and puts, and recalling that from (18), the standard error of  $\hat{\Pi}^S$  is given by

$$s.e.(\hat{\Pi}^S) = s.e.(\hat{B} - \hat{\mu}^\dagger) = \left\{ \omega^2 s.e.(\hat{B}_c - \hat{\mu}_c^\dagger)^2 + (1 - \omega)^2 s.e.(\hat{B}_p - \hat{\mu}_p^\dagger)^2 \right\}^{1/2}. \quad (21)$$

Recall that the point estimate of asset price bubble is given by (18):  $\hat{\Pi}^S = S + \hat{B} - \hat{\mu}^\dagger$ . Note that it is not consistent for  $\Pi^S$  as the fundamental asset value is not point-identified (due to the presence of term  $A$ ). Nonetheless, it is partially identified and there exists a consistent interval estimator for  $\Pi^S$ , which we turn to next.

Since  $\hat{\mu}^\dagger$  is asymptotically normal by Proposition 3, and since  $\hat{\mu}$  is (asymptotically) a linear function of  $\hat{\mu}^\dagger$ , we may construct the confidence interval for the fundamental asset value (and hence the bubble) by normal approximation. Fix the confidence level at  $100(1 - \alpha)\%$ . The minimum possible fundamental asset value has confidence interval given by the bounds  $\hat{\mu}^\dagger - (\hat{B} - A_\ell) \mp z_{\alpha/2} \cdot s.e.(\hat{\Pi}^S)$ , so that the confidence interval for the maximum possible bubble is  $S - \hat{\mu}^\dagger + (\hat{B} - A_\ell) \mp z_{\alpha/2} \cdot s.e.(\hat{\Pi}^S)$ . Similarly, the confidence interval for the minimum possible bubble is  $S - \hat{\mu}^\dagger + (\hat{B} - A_u) \mp z_{\alpha/2} \cdot s.e.(\hat{\Pi}^S)$ . The  $100(1 - \alpha)\%$  confidence interval of asset price bubble is the union of all

the confidence intervals associated with all (partially identified) bubbles, i.e.,

$$[S - \hat{\mu}^\dagger + (\hat{B} - A_u) - z_{\alpha/2} \cdot s.e.(\hat{\Pi}^S), S - \hat{\mu}^\dagger + (\hat{B} - A_\ell) + z_{\alpha/2} \cdot s.e.(\hat{\Pi}^S)],$$

as given in (19).

## A2 Regularity Conditions of Proposition 3

This section states the regularity conditions in Proposition 3. Some of the regularity conditions are adaptation of Condition 1 in Section 6.6.2 of Fan and Yao (2003).

- (a) The option prices (as a function of strike  $k$ ) are given by  $C(k) = m_c(k) + \epsilon_c(k)$  and  $P(k) = m_p(k) + \epsilon_p(k)$ , where  $m_c(k)$  and  $m_p(k)$  are finite, strictly positive, and twice continuously differentiable at finite  $k > 0$ .
- (b) The error  $\epsilon_c(k)$  ( $\epsilon_p(k)$ ) has mean zero and variance  $0 < \sigma_{\epsilon_c}^2(k) < \infty$  ( $\sigma_{\epsilon_p}^2(k) < \infty$ ). Both  $\sigma_{\epsilon_c}^2(k)$  and  $\sigma_{\epsilon_p}^2(k)$  are continuous in  $k$ .
- (c) The error process  $\{\epsilon_c(k)\}_k$  (similarly for  $\{\epsilon_p(k)\}_k$ ) is either  $\rho$ -mixing with  $\sum_{j=1}^{\infty} \rho(j) < \infty$ , or  $\alpha$ -mixing with  $\sum_{j=1}^{\infty} j^a [\alpha(j)]^{1-2/\delta} < \infty$  for some  $\delta > 2$  and  $a > 1 - 2/\delta$ .
- (d) There exists a sequence of positive integers satisfying  $r_n \rightarrow \infty$  and  $r_n = o(\sqrt{nh_n})$  such that  $\sqrt{n/h_n} \rho(r_n) \rightarrow 0$  and  $\sqrt{n/h_n} \alpha(r_n) \rightarrow 0$  as  $n \rightarrow \infty$ .
- (e)  $\epsilon_c(k)$  and  $\epsilon_p(k')$  are independent for all  $k, k'$ .
- (f) The kernel function in local polynomial estimation is bounded with bounded support.

*Remark:* Condition (a) sets the stage for implementing local quadratic estimation, which yields the estimates for  $m_c^{(1)}(k)$  and  $m_p^{(1)}(k)$ . Conditions (b)-(d) exert control on the pricing errors in the cross section (heteroskedastic and weakly dependent errors are allowed), whereas condition (e) assumes that the pricing errors of puts and calls are independent.<sup>29</sup> Condition (f) is a technical restriction assumed in Fan and Yao (2003), but it can be relaxed to include light-tailed kernels (e.g., Gaussian kernel).

---

<sup>29</sup>The error independence between calls and puts offers convenience while computing  $s.e.(\hat{\Pi}^S)$  in (21) when  $\omega \neq 0$  or 1. This restriction can be relaxed, although it is not clear how to model the dependence of the call and put errors in an empirically relevant way.

### A3 Technical Proofs

In all the proofs, we suppress the time subscript  $t$  and the time-to-maturity  $\tau$ , both of which are taken as given, unless it is necessary to avoid confusion.

#### A3.1 Proof of Proposition 1

See Theorems 4.4 and 6.4 of Jarrow, Protter, and Shimbo (2010). Note that no dominance is not needed to obtain the result.

#### A3.2 Proof of Lemma 1

First we show that  $C^*(k)$  and  $P^*(k)$  are differentiable. Denote  $1_{\{\cdot\}}$  the indicator function, and  $\delta_{\{\cdot\}}$  the Dirac delta function. Now rewrite  $C^*(k) = \int_0^\infty e^{-r\tau} 1_{\{s \geq k\}}(s-k) dQ^*(s)$  and define the integrand  $h(k) = e^{-r\tau} 1_{\{s \geq k\}}(s-k)$ . Its derivative is  $h'(k) = -e^{-r\tau} 1_{\{s \geq k\}} - e^{-r\tau} \delta_{\{s=k\}}(s-k) = -e^{-r\tau} 1_{\{s \geq k\}}$ , which is bounded as  $|h'(k)| \leq e^{-r\tau} \leq 1$ . The upper bound is integrable as  $\int_0^\infty dQ^*(s) = 1$ . It follows from the dominated convergence theorem that  $C^*(k)$  is differentiable. Similar argument shows that  $P^*(k)$  is differentiable.

Differentiating both sides of (4) with respect to  $k$  yields

$$\begin{aligned} \frac{\partial C(k)}{\partial k} &= \frac{\partial C^*(k)}{\partial k} + \frac{\partial \Pi^C(k)}{\partial k} \\ &= \frac{\partial}{\partial k} \int_k^\infty e^{-r\tau}(s-k) dQ^*(s) + \frac{\partial \Pi^C(k)}{\partial k} \\ &= - \int_k^\infty e^{-r\tau} dQ^*(s) + \frac{\partial \Pi^C(k)}{\partial k} \\ &= -e^{-r\tau}[1 - Q^*(k)] + \frac{\partial \Pi^C(k)}{\partial k}, \end{aligned}$$

where the last step follows from integration by parts. We therefore obtain the first equality (6). The second equality (7) is obtained by a similar argument applied to (5). The proof is now completed.

#### A3.3 Proof of Lemma 2

The proof of Lemma 1 shows that

$$\frac{\partial C^*(k)}{\partial k} = -e^{-r\tau}[1 - Q^*(k)].$$

The first equality in the lemma is obtained by straightforward integration of the above equality over  $k \in (u, \infty)$  and by noting that  $C^*(\infty) = 0$ . The second equality can be shown analogously.

### A3.4 Proof of Proposition 2

To prepare for the proof, we first note that the fundamental values of options are non-negative by definition, and that the option bubbles are non-negative by Proposition 1. This yields

$$0 \leq C^*(u_c) \leq C(u_c) \quad (22)$$

and

$$0 \leq P^*(\ell_p) = P(\ell_p). \quad (23)$$

(a) We evaluate the fundamental asset price by breaking down the integration into five parts:

$$\begin{aligned} \mu^* &= e^{-r\tau} \left( \int_0^{\ell_p} + \int_{\ell_p}^{\ell_c} + \int_{\ell_c}^{u_p} + \int_{u_p}^{u_c} + \int_{u_c}^{\infty} \right) [1 - Q^*(s)] ds \\ &= e^{-r\tau} \left\{ \ell_p + \left( \int_{\ell_p}^{\ell_c} + \int_{\ell_c}^{u_p} + \int_{u_p}^{u_c} \right) [1 - Q^*(s)] ds \right\} + C^*(u_c) - P^*(\ell_p), \end{aligned} \quad (24)$$

where the last step is by Lemma 2. The integrals in the last line are computed in turn. Since only put data are available in  $[\ell_p, \ell_c]$ , we apply (7) to the integrand, yielding

$$\begin{aligned} e^{-r\tau} \int_{\ell_p}^{\ell_c} [1 - Q^*(s)] ds &= e^{-r\tau} (\ell_c - \ell_p) - \int_{\ell_p}^{\ell_c} \frac{\partial P(s)}{\partial s} ds \\ &= e^{-r\tau} (\ell_c - \ell_p) + P(\ell_p) - P(\ell_c). \end{aligned}$$

Since only call data are available in  $[u_p, u_c]$ , we apply (6) to the integrand, yielding

$$e^{-r\tau} \int_{u_p}^{u_c} [1 - Q^*(s)] ds = - \int_{u_p}^{u_c} \frac{\partial C(s)}{\partial s} ds = C(u_p) - C(u_c).$$

For the middle interval  $[\ell_c, u_p]$ , both call and put data are available, so we may apply (6) or (7) to

the integrand, yielding

$$e^{-r\tau} \int_{\ell_c}^{u_p} [1 - Q^*(s)] ds = - \int_{\ell_c}^{u_p} \frac{\partial C(s)}{\partial s} ds = C(\ell_c) - C(u_p)$$

or

$$e^{-r\tau} \int_{\ell_c}^{u_p} [1 - Q^*(s)] ds = e^{-r\tau} (u_p - \ell_c) - \int_{\ell_c}^{u_p} \frac{\partial P(s)}{\partial s} ds = e^{-r\tau} (u_p - \ell_c) + P(\ell_c) - P(u_p).$$

Substituting into (24), we obtain

$$\begin{aligned} \mu^* &= e^{-r\tau} \ell_c + [P(\ell_p) - P(\ell_c)] + [C(\ell_c) - C(u_p)] + [C(u_p) - C(u_c)] + C^*(u_c) - P^*(\ell_p) \\ &= m_1 + C^*(u_c) - P^*(\ell_p), \end{aligned}$$

where  $m_1 := e^{-r\tau} \ell_c + C(\ell_c) - C(u_c) - P(\ell_c) + P(\ell_p)$ , or

$$\begin{aligned} \mu^* &= e^{-r\tau} u_p + [P(\ell_p) - P(\ell_c)] + [P(\ell_c) - P(u_p)] + [C(u_p) - C(u_c)] + C^*(u_c) - P^*(\ell_p) \\ &= m_2 + C^*(u_c) - P^*(\ell_p), \end{aligned}$$

where  $m_2 := e^{-r\tau} u_p + C(u_p) - C(u_c) + P(\ell_p) - P(u_p)$ . Consequently, by Lemma 2, the bounds for asset price bubble are thus deduced as

$$S_t - m_j - C(u_c) \leq \Pi_t^S(\tau) \leq S_t - m_j + P(\ell_p), \quad (25)$$

for  $j = 1, 2$ , or any linear combination of the two inequalities.

(b) The fundamental asset price is given by:

$$\mu^* = e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} + C^*(u_c) - P^*(\ell_p).$$

We may evaluate the integral using calls. The relation (6) from Lemma 1 implies that

$$e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} = e^{-r\tau} \ell_p - \int_{\ell_p}^{u_c} \frac{\partial C_t(s)}{\partial k} ds = e^{-r\tau} \ell_p - C(u_c) + C(\ell_p) =: m_1.$$

Alternatively, we may evaluate the integral using puts. Using (7) from Lemma 1, we have

$$e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} = e^{-r\tau} u_c - \int_{\ell_p}^{u_c} \frac{\partial P_t(s)}{\partial k} ds = e^{-r\tau} u_c - P(u_c) + P(\ell_p) =: m_2.$$

By the same argument as in part (a), the bounds for asset price bubble are given by (25) with  $m_j$  defined above.

(c) The fundamental asset price is given by:

$$\mu^* = e^{-r\tau} \left\{ \ell_p + \left( \int_{\ell_p}^{\ell_c} + \int_{\ell_c}^{u_c} \right) [1 - Q^*(s)] ds \right\} + C^*(u_c) - P^*(\ell_p).$$

The first integral is evaluated using puts only, while the second one can be evaluated using calls or puts. With the help of (6) and (7), we obtain

$$\begin{aligned} e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} &= e^{-r\tau} \ell_c - \int_{\ell_p}^{\ell_c} \frac{\partial P_t(s)}{\partial k} ds - \int_{\ell_c}^{u_c} \frac{\partial C_t(s)}{\partial k} ds \\ &= e^{-r\tau} \ell_c + P(\ell_p) - P(\ell_c) + C(\ell_c) - C(u_c) =: m_1, \end{aligned}$$

or

$$e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} = e^{-r\tau} u_c - \int_{\ell_p}^{u_c} \frac{\partial P_t(s)}{\partial k} ds = e^{-r\tau} u_c + P(\ell_p) - P(u_c) =: m_2.$$

By the same argument as in part (a), the bounds for asset price bubble are given by (25) with  $m_j$  defined above.

(d) We compute the fundamental asset price as follows:

$$\mu^* = e^{-r\tau} \left\{ \ell_p + \left( \int_{\ell_p}^{u_p} + \int_{u_p}^{u_c} \right) [1 - Q^*(s)] ds \right\} + C^*(u_c) - P^*(\ell_p).$$

The first integral can be evaluated using puts calls or puts, while the second one is evaluated using calls only. With the help of (6) and (7), we obtain

$$e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} = e^{-r\tau} \ell_p - \int_{\ell_p}^{u_c} \frac{\partial C_t(s)}{\partial k} ds = e^{-r\tau} \ell_p + C(\ell_p) - C(u_c) =: m_1,$$

or

$$\begin{aligned} e^{-r\tau} \left\{ \ell_p + \int_{\ell_p}^{u_c} [1 - Q^*(s)] ds \right\} &= e^{-r\tau} u_p - \int_{\ell_p}^{u_p} \frac{\partial P_t(s)}{\partial k} ds - \int_{u_p}^{u_c} \frac{\partial C_t(s)}{\partial k} ds \\ &= e^{-r\tau} u_p + P(\ell_p) - P(u_p) + C(u_p) - C(u_c) =: m_2. \end{aligned}$$

By the same argument as in part (a), the bounds for asset price bubble are given by (25) with  $m_j$  defined above.

(e) The fundamental asset price is given by:

$$\mu^* = e^{-r\tau} \left\{ \ell_p + \left( \int_{\ell_p}^{u_p} + \int_{u_p}^{\ell_c} + \int_{\ell_c}^{u_c} \right) [1 - Q^*(s)] ds \right\} + C^*(u_c) - P^*(\ell_p).$$

The first integral is evaluated by puts only, and the third integral is evaluated by calls only:

$$\begin{aligned} \int_{\ell_p}^{u_p} [1 - Q^*(s)] ds &= e^{-r\tau} (u_p - \ell_p) - \int_{\ell_p}^{u_p} \frac{\partial P_t(s)}{\partial k} ds = e^{-r\tau} (u_p - \ell_p) + P(\ell_p) - P(u_p), \\ \int_{\ell_c}^{u_c} [1 - Q^*(s)] ds &= - \int_{\ell_c}^{u_c} \frac{\partial C_t(s)}{\partial k} ds = C(\ell_c) - C(u_c). \end{aligned}$$

No data are available for computing the middle integral, but we know that the following bounds are valid:

$$e^{-r\tau} \int_{u_p}^{\ell_c} [1 - Q^*(s)] ds \in [0, e^{-r\tau} (\ell_c - u_p)].$$

Combining the above expressions, and applying Lemma 2, we can bound the fundamental price as follows:

$$e^{-r\tau} u_p - P(u_p) + C(\ell_c) - C(u_c) \leq \mu^* \leq e^{-r\tau} \ell_c + P(\ell_p) - P(u_p) + C(\ell_c).$$

The stated bounds for asset price bubble follow immediately.

(f) The fundamental asset price can be expressed as:

$$\begin{aligned} \mu^* &= e^{-r\tau} \left\{ \left( \int_{-\infty}^{\ell_p} + \int_{u_c}^{\infty} - \int_{u_c}^{\ell_p} \right) [1 - Q^*(s)] ds \right\} \\ &= e^{-r\tau} \ell_p - P^*(\ell_p) + C^*(u_c) - e^{-r\tau} \int_{u_c}^{\ell_p} [1 - Q^*(s)] ds. \end{aligned}$$



No data are available for computing the integral on the last line, but we know that the following bounds are valid:

$$e^{-r\tau} \int_{u_c}^{\ell_p} [1 - Q^*(s)] ds \in [0, e^{-r\tau}(\ell_p - u_c)].$$

By Lemma 2, we can bound the fundamental price as follows:

$$e^{-r\tau} u_c - P(\ell_p) \leq \mu^* \leq e^{-r\tau} \ell_p + C(u_c).$$

The stated bounds for asset price bubble follow immediately.

### A3.5 Proof of Corollary 1

We simply observe that, under ND, put-call parity in market prices hold, and the inequalities in cases (a)-(d) reduce to the stated bound.

### A3.6 Proof of Lemma 3

By (12) and (13), we have

$$\mu^\dagger = e^{-r\tau} \left\{ \ell + \int_{\ell}^u \left[ 1 - \frac{Q^*(s) - Q^*(\ell)}{Q^*(u) - Q^*(\ell)} \right] ds \right\}.$$

On the other hand, we deduce from (8) that

$$\mu^* = e^{-r\tau} \left\{ \ell + \int_{\ell}^u [1 - Q^*(s)] ds \right\} + C^*(u) - P^*(\ell).$$

By subtraction, we obtain

$$\begin{aligned} \mu^\dagger - \mu^* &= e^{-r\tau} \int_{\ell}^u \left[ Q^*(s) - \frac{Q^*(s) - Q^*(\ell)}{Q^*(u) - Q^*(\ell)} \right] ds + P^*(\ell) - C^*(u) \\ &= \left[ \frac{e^{-r\tau} Q^*(\ell)(u - \ell)}{Q^*(u) - Q^*(\ell)} - \left( \frac{1}{Q^*(u) - Q^*(\ell)} - 1 \right) e^{-r\tau} \int_{\ell}^u Q^*(s) ds \right] - [C^*(u) - P^*(\ell)]. \\ &=: B - A. \end{aligned}$$

To obtain the bounds for  $A$ , we note that

$$\begin{aligned} 0 &\leq C^*(u) \leq C(u), \\ 0 &\leq P^*(\ell) = P(\ell), \end{aligned}$$

where the lower bounds are by the non-negative nature of the fundamental option values, while the upper bounds are obtained from the definitions of option bubbles (4)-(5) and Proposition 1. The stated expressions for  $B$  are obtained by integrating (6) and (7) and applying Lemma 2 and Assumption INK. The proof is now completed.

### A3.7 Proof of Proposition 3

From (14), we see that  $\hat{Q}^*(s)$  is linear in  $\hat{m}^{(1)}(s)$ , and so it inherits the pointwise convergence and asymptotic normality of  $\hat{m}^{(1)}(s)$  (Theorem 6.3 of Fan and Yao (2005), with  $\nu = 1$  and  $p = 2$ ). By pointwise convergence of  $\hat{Q}^*(s)$  for each  $s \in [\ell, u]$ , we deduce from (12) that

$$\hat{Q}^\dagger(s) = \frac{\hat{Q}^*(s) - \hat{Q}^*(\ell)}{\hat{Q}^*(u) - \hat{Q}^*(\ell)} = \frac{\hat{Q}^*(s) - Q^*(\ell)}{Q^*(u) - Q^*(\ell)} + o_p(1).$$

It follows from linearity that pointwise convergence and asymptotic normality hold for  $\hat{Q}^\dagger(s)$ . From the definition of  $\hat{\mu}^\dagger$  in (17), we have

$$\begin{aligned} \hat{\mu}^\dagger &= e^{-r\tau} \ell + e^{-r\tau} \sum_{i=1}^{n_s} [1 - \hat{Q}^\dagger(s_i)] \Delta s_i \\ &= \left\{ e^{-r\tau} \ell + e^{-r\tau} \sum_{i=1}^{n_s} [1 - Q^\dagger(s_i)] \Delta s_i \right\} - e^{-r\tau} \sum_{i=1}^{n_s} [\hat{Q}^\dagger(s_i) - Q^\dagger(s_i)] \Delta s_i \\ &=: (I) + (II). \end{aligned} \tag{26}$$

The first term (I) differs from  $\mu^\dagger$  by  $O(n_s^{-1})$ . Indeed, we note that

$$\begin{aligned}
(I) - \mu^\dagger &= e^{-r\tau} \sum_{i=1}^{n_s} \left[ \int_{s_{i-1}}^{s_i} Q^\dagger(s) ds - Q^\dagger(s_i) \Delta s_i \right] \\
&\leq e^{-r\tau} \sum_{i=1}^{n_s} \left[ \sup_{s \in [s_{i-1}, s_i]} Q^\dagger(s) - Q^\dagger(s_i) \right] \Delta s_i \\
&\leq e^{-r\tau} \sum_{i=1}^{n_s} a_i (\Delta s_i)^2 = O(n_s^{-1}),
\end{aligned}$$

where the second inequality is obtained by the continuity of  $Q^\dagger$ . The second term (II) has an asymptotic bias of order  $O(h^3)$  (Fan and Yao (2005), Theorem 6.3). The bias of  $\hat{\mu}^\dagger$  is thus of order  $O(n_s^{-1} + h^3)$ . The variance of  $\hat{\mu}^\dagger$  is contributed by (II), and is equal to

$$\begin{aligned}
Var(\hat{\mu}^\dagger) &= e^{-2r\tau} Var \left( \sum_{i=1}^{n_s} [\hat{Q}^\dagger(s_i) - Q^\dagger(s_i)] \Delta s_i \right) \\
&= Var \left( \sum_{i=1}^{n_s} \hat{m}^{(1)}(s_i) \Delta s_i \right) \\
&= Var \left( \frac{1}{n_s} \sum_{i=1}^{n_s} \hat{m}^{(1)}(s_i) \right), \tag{27}
\end{aligned}$$

It has a uniform upper bound:

$$Var(\hat{\mu}^\dagger) \leq Var \left( \sup_{i=1, \dots, n_s} \hat{m}^{(1)}(s_i) \right) = O \left( \frac{\log(n)}{nh^3} \right),$$

which holds uniformly in  $n_s$  (Fan and Yao (2005), Theorem 6.5). It follows that  $\hat{\mu}^\dagger$  is consistent for  $\mu^\dagger$  as  $bias(\hat{\mu}^\dagger) = O(n_s^{-1} + h^3) = o(1)$  and  $Var(\hat{\mu}^\dagger) = O \left( \frac{\log(n)}{nh^3} \right) = o(1)$  by the assumed asymptotic rates on  $n_s$ ,  $n$  and  $h$ .

Subtracting (I) and the bias from both sides of (26) and then multiplying by  $\sqrt{nh^3}$ , we obtain

$$\sqrt{nh^3}(\hat{\mu}^\dagger - \mu^\dagger - bias(\hat{\mu}^\dagger)) = -\sqrt{nh^3} e^{-r\tau} \sum_{i=1}^{n_s} [\hat{Q}^\dagger(s_i) - Q^\dagger(s_i)] \Delta s_i.$$

By (27), the variance of the right-hand-side is

$$\begin{aligned} V_{n_s, n, h} &= nh^3 \text{Var} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} \hat{m}^{(1)}(s_i) \right) \\ &= \frac{nh^3}{n_s} \frac{1}{n_s} \sum_{i=1}^{n_s} \text{Var}[\hat{m}^{(1)}(s_i)] + \frac{nh^3}{n_s} \frac{1}{n_s} \sum_{\ell=1}^{n_s-1} \sum_{i=1}^{n_s-1} \text{Cov}[\hat{m}^{(1)}(s_i), \hat{m}^{(1)}(s_{i+\ell})]. \end{aligned}$$

The first term on the last line is  $\frac{nh^3}{n_s} O\left(\frac{\log(n)}{nh^3}\right) = O\left(\frac{\log(n)}{n_s}\right)$ , which follows the uniform bound:  $\text{Var}[\hat{m}^{(1)}(s)] = O\left(\frac{\log(n)}{nh^3}\right)$  uniformly in  $s$  (Fan and Yao (2005), Theorem 6.5). The second term is  $O\left(\frac{\log(n)}{n_s}\right)$  by the absolute summability of the covariances and the regularity conditions on the underlying mixing process. We thus deduce that  $V_{n_s, n, h} = O(1)$ , as  $\frac{\log(n)}{n_s} = O(1)$  by assumption. Finally, the asymptotic normality of  $\hat{\mu}^\dagger$  is inherited from that of  $\hat{Q}^\dagger(s)$  as  $\hat{\mu}^\dagger$  is linear in  $\hat{Q}^\dagger(s)$ . The proof is now completed.

### A3.8 Proof of Lemma A1

We first consider using only put price data to evaluate  $\hat{\mu}$ . By simple algebra, we obtain

$$\begin{aligned} \hat{B} - B &= \frac{e^{-r\tau}(u - \ell)Q^*(u) - [P(u) - P(\ell)]}{[\hat{Q}^*(u) - \hat{Q}^*(\ell)][Q^*(u) - Q^*(\ell)]} [\hat{Q}^*(\ell) - Q^*(\ell)] \\ &\quad + \frac{[P(u) - P(\ell)] - e^{-r\tau}(u - \ell)Q^*(\ell)}{[\hat{Q}^*(u) - \hat{Q}^*(\ell)][Q^*(u) - Q^*(\ell)]} [\hat{Q}^*(u) - Q^*(u)], \end{aligned}$$

and

$$\begin{aligned} \hat{\mu}^\dagger - \mu^\dagger &= e^{-r\tau} \frac{\hat{Q}^*(u)(u - \ell) - \sum_{i=1}^{n_s} Q^*(s_i) \Delta s_i}{[\hat{Q}^*(u) - \hat{Q}^*(\ell)][Q^*(u) - Q^*(\ell)]} [\hat{Q}^*(\ell) - Q^*(\ell)] \\ &\quad - e^{-r\tau} \sum_{i=1}^{n_s} \frac{\Delta s_i}{\hat{Q}^*(u) - \hat{Q}^*(\ell)} [\hat{Q}^*(s_i) - Q^*(s_i)] \\ &\quad + e^{-r\tau} \frac{\sum_{i=1}^{n_s} Q^*(s_i) \Delta s_i - \hat{Q}^*(\ell)(u - \ell)}{[\hat{Q}^*(u) - \hat{Q}^*(\ell)][Q^*(u) - Q^*(\ell)]} [\hat{Q}^*(u) - Q^*(u)]. \end{aligned}$$

It follows that

$$(\hat{B} - \hat{\mu}^\dagger) - (B - \mu^\dagger) = e^{-r\tau} \sum_{i=1}^{n_s} w_i [\hat{Q}^*(s_i) - Q^*(s_i)],$$

where the weights are as given in the statement of the lemma. The result for using only call price data is obtained by replacing  $P(u) - P(\ell)$  by  $C(u) - C(\ell) + e^{-r\tau}(u - \ell)$ .

# Appendix Figures

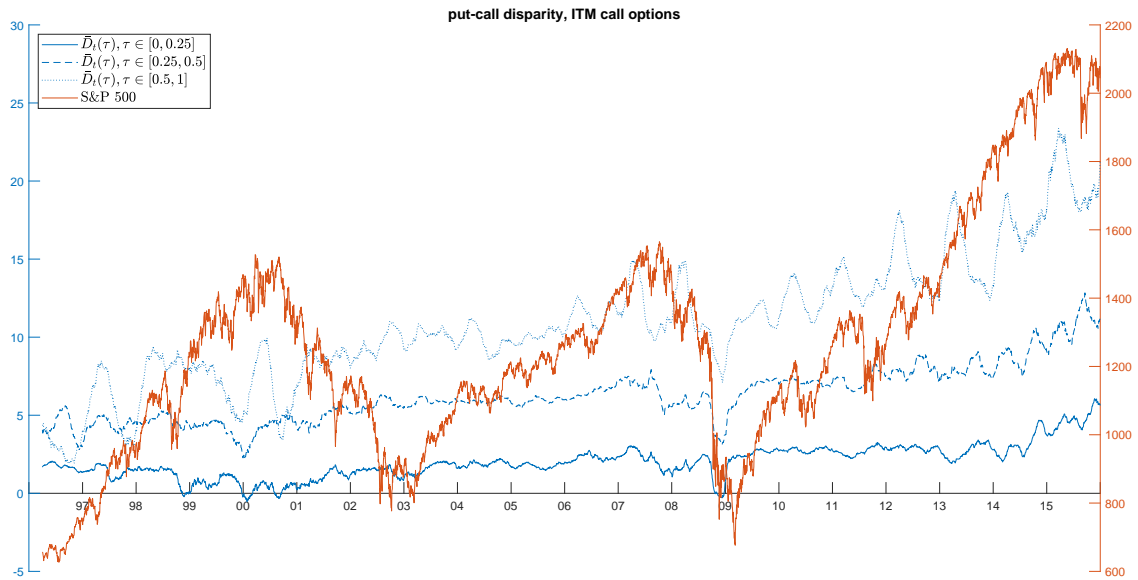


Figure A1: Put-call disparity averaged over strikes associated with ITM calls/OTM puts (log-moneyness  $< -0.1$ ).

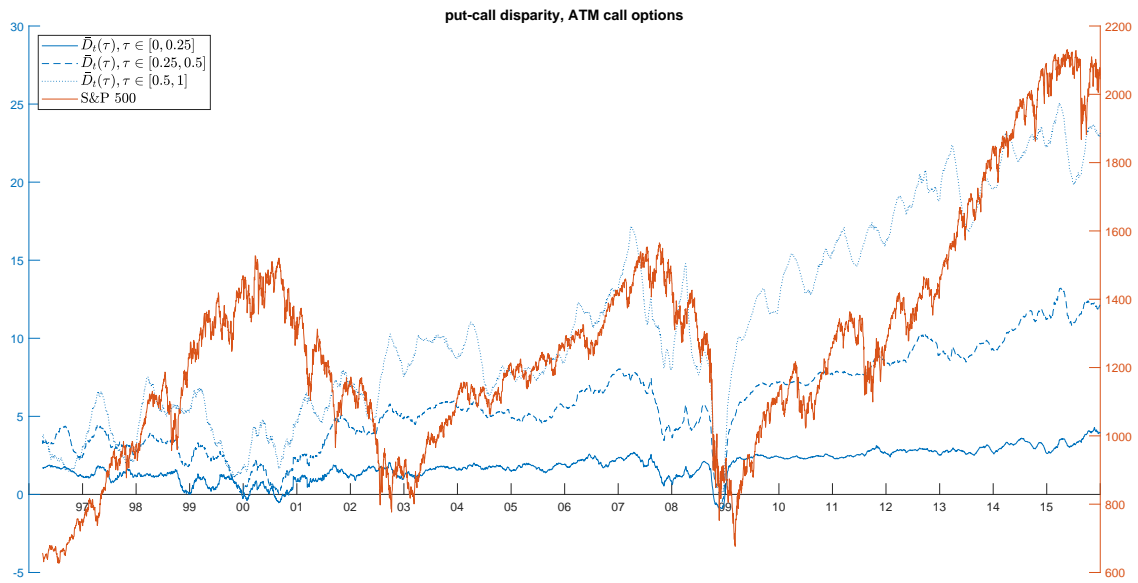


Figure A2: Put-call disparity averaged over strikes associated with ATM calls/puts ( $-0.1 < \text{log-moneyness} < 0.1$ ).

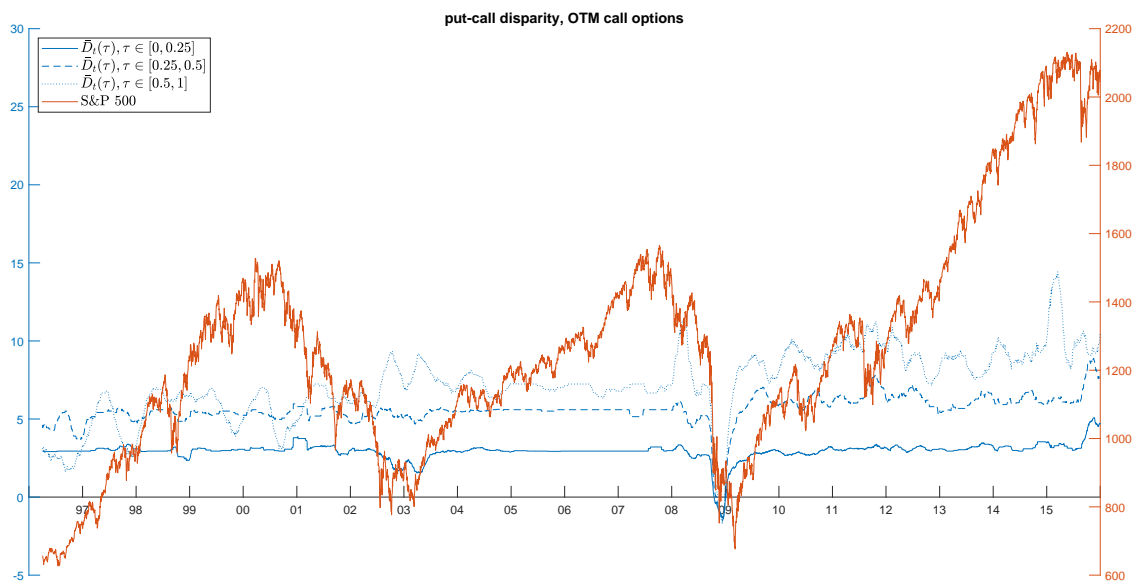


Figure A3: Put-call disparity averaged over strikes associated with OTM calls/ITM puts (log-moneyness  $> 0.1$ ).