# Forecasting Volatility in the Presence of Limits to Arbitrage

June 7, 2013

Lu Hong, Loyola University Chicago, 1 East Pearson St., Chicago, IL 60611; lhong@luc.edu, (312) 915-7067

Tom Nohel*, Loyola University Chicago, 1 East Pearson St., Chicago, IL 60611; tnohel@luc.edu, (312) 915-7065

Steven Todd, Loyola University Chicago, 1 East Pearson St., Chicago, IL 60611; stodd@luc.edu, (312) 915-7218

## Abstract

In this paper, we develop a novel model to forecast the volatility of S&P 500 futures returns by considering measures of limits to arbitrage. When arbitrageurs face constraints on their trading strategies, option prices can become disconnected from fundamentals, resulting in a premium that reflects the limits to arbitrage. The corresponding market based implied volatility may therefore also contain these distortions. We argue that limits to arbitrage can be systematic or idiosyncratic and we search for proxies to capture these effects. Our contributions are both conceptual and empirical. Conceptually, the distinction between systematic and idiosyncratic effects of limits to arbitrage can shed light on relative asset prices as exemplified by this particular study. Empirically, our volatility forecasting model explains 71% of the variation in realized volatility, a substantial improvement over a naive forecast based only on lagged realized volatility, which produces an $R^2$ of 53%.

* Corresponding author

## Forecasting Volatility in the Presence of Limits to Arbitrage

### 1. Introduction

Forecasting volatility is an essential task for many financial market participants. The level and evolution of volatility directly impact hedge funds, that specialize in volatility trading strategies, financial services firms, that provide insurance against volatility, regulators, who seek to stabilize markets, and risk managers, who measure and manage a portfolio's value-at-risk (VAR). To the extent that realized and anticipated volatility affect current and future asset prices, anyone who manages money can benefit from a better forecast of volatility.

Volatility is an essential parameter used to price option contracts. Alternatively, given a set of options with known strike prices and expiration dates, and given the risk-free rate of interest, the observed market prices of options can be used to back out an estimate of the volatility expected to prevail over the life of the option, i.e., the *implied volatility*. Assuming frictionless markets, and assuming the underlying stock price process is characterized by geometric Brownian motion, the implied volatility is an unbiased and efficient estimate of the volatility that will prevail over the life of the option because potential arbitrageurs face no impediments to capitalizing on any arbitrage opportunities that arise and as a result will act to keep prices in line.

However, in the presence of transaction costs, financing constraints, and other *limits to arbitrage*, option prices can become distorted, since potential arbitrageurs will scale back their activities due to a lack of capital, significant transaction costs, or a perception that the risk-reward trade-off is not favorable. When option prices are distorted, the implied volatility from an option pricing model is also distorted, imparting noise and/or bias into volatility forecasts. In fact, prior studies have consistently shown that implied volatility is a biased forecast of future realized volatility (see Christensen and Prabhala, 1998; Jiang and Tian, 2005; and Poteshman, 2000, among others). In this paper, we look to extract the premium from option prices that reflects limits to arbitrage, with the intent of improving the forecasting power of implied volatility.

Recently, the finance literature has focused a great deal of attention on the concept of liquidity – the ease and speed with which investors can enter and exit positions at reasonable prices. A lack of liquidity can lead to temporary price discrepancies between similar baskets of assets. Arbitrageurs seek to capitalize on such discrepancies by executing convergence trades, providing liquidity where it is lacking and profiting as prices correct. At Salomon Brothers, John Merriwether mastered the art of convergence trading, focusing on simple pricing discrepancies between on-the-run and off-the-run Treasury securities. Other hedge funds or proprietary trading desks seek to take advantage of more complex pricing anomalies, such as those between closed-end funds and the assets that they invest in.

Garleanu, Pedersen, and Poteshman (2009) show that market-makers (and by extension, arbitrageurs) who usually maintain net short positions in equity index options, can profit handsomely by selling insurance to end-users who are hedging portfolio risks. When end-user demand for insurance is high, option writers require greater compensation as providers of liquidity. As a result, option prices rise, which in turn, begets an increase in implied volatility. Clearly, asset prices, liquidity and limits to arbitrage are very much inter-connected.

Brunnermeier and Pedersen (2009) make the crucial observation that arbitrageurs typically invest outside capital, relying on so-called *funding liquidity* – the ease and speed with which traders/arbitrageurs can replace withdrawn capital. Arbitrageurs may face different limits to arbitrage in different markets, but all arbitrageurs are exposed to the possibility of funding liquidity shocks. A sudden drop in funding liquidity can turn liquidity providers (i.e., arbitrageurs) into liquidity consumers, with adverse consequences for the markets in which the arbitrageurs operate. Hence, funding liquidity and market liquidity are inextricably linked. Brunnermeier and Pedersen (2009) develop a model and show how funding liquidity and market liquidity can work in tandem to create illiquidity spirals. Some recent papers confirm that funding liquidity and market liquidity are connected.[1]

Hu, Pan, and Wang (2012), hereafter HPW, propose a market-wide liquidity measure based on price deviations in the U.S. Treasury market, the world's most liquid market and a market whose securities most closely approximate risk-free securities. When there is ample arbitrage capital available, all Treasury securities are anchored to the zero-coupon yield curve. However, when arbitrage capital is in short supply, very-liquid on-the-run Treasury securities trade at a premium, relative to their off-the-run breathren. HPW average these Treasury security price differences across a wide range of maturities to produce a measure of market-wide liquidity. When the price disparities between on-the-run and off-the-run Treasury securities are high, funding capital must be

[1] See, for example, Mitchell and Pulvino (2012), Brunnermeier and Pedersen (2009), and Mitchell, Petersen, and Pulvino (2007). Related papers include Adrian and Shin (2010), Gorton and Metrick (2010), and Duffie (2010). Also, see Deuskar (2006) for a model wherein volatility begets illiquidity, leading to more volatility.

in short-supply, because the obvious set of convergence trades required by arbitrageurs to correct relative mispricing in the Treasury market are not being made.

Similarly, when a closed-end fund (CEF) trades at a discount or premium to its net asset value (NAV), arbitrageurs should step in to force a convergence in prices. That they do not suggests that there are limits to arbitrage, either in the form of unavailable capital, or capital that is very expensive. Pontiff (1996) argues that CEF arbitrage is costly. Not only are arbitrageurs exposed to the risk that outside capital will flee, but there remain significant, unhedgeable risks (e.g., basis risk) due to the fact that CEF portfolios are observable only at a quarterly frequency with a 45-day delay. Moreover, trading costs (e.g., taking short positions in CEF shares or their portfolio holdings) can be prohibitive, especially for CEFs that invest in illiquid assets (see Nohel, Todd, and Wang, 2013).

Pontiff (1996) shows that CEF deviations from NAV are an increasing function of arbitrage costs. In the presence of high funding or transaction costs, arbitrageurs will require larger rewards for their activities, resulting in ever larger deviations from fundamental value. We argue that in stressed markets, characterized by high volatility, the costs of arbitrage increase and we should expect larger deviations from fundamental value. We acknowledge two distinct types of costs for arbitrageurs: costs resulting from systematic liquidity constraints or liquidity shocks; and costs associated with idiosyncratic illiquidity or event risks specific to a particular asset or portfolio (such as changes in margin requirements, restrictions on shorting, costs related to basis risk, uncertainty about portfolio holdings, or unhedgeable risks, such as higher moments of risk due to market incompleteness).

One might think that S&P 500 index option traders need not worry about idiosyncratic liquidity because the S&P 500 is a diversified basket and there is no basis risk between index options and S&P 500 futures. However, the largest arbitrageurs (e.g., hedge funds and money-center banks) operate in several asset classes/markets, and an idiosyncratic shock in one market can quickly reverberate into other markets. Hence, both systematic and idiosyncratic liquidity are likely to converge during a financial crisis.[2] Moreover, arbitrage trades that aren't unwound intra-day are exposed to higher moments of risk, related to stochastic volatility.

Our conjecture is that when arbitrage capital is in short supply, end-user demand for index options is high, resulting in over-priced options (commensurately overstated implied volatilities) and an upwardly biased and inefficient forecast of future realized volatility. We construct an index that measures the aggregate mispricing among equity closed-end funds. Whereas the noise measure of HPW captures systematic liquidity constraints, we argue that our measure captures asset-specific illiquidity, such as event risks, basis risks, and unhedgeable risks, in addition to the risk of systematic liquidity shocks.

We construct a dataset of S&P 500 futures and futures options spanning the period 1997 - 2008 to test our volatility forecasting model. We use futures prices in 5-minute intervals to construct a realized volatility series. We sample end-of-day option prices each month, focusing on next-to-expire contracts (with just under four weeks until

[2] Mitchell and Pulvino (2012) show that during the recent financial crisis, a reduction in lending by prime brokers created tremendous opportunities for hedge funds engaging in convertible arbitrage and credit arbitrage (CDS versus bonds). Such opportunities persisted for months, suggesting that arbitrage capital was in short supply. Moreover, contemporaneous and subsequent hedge fund de-leveraging had a lingering impact on merger spreads and CEF discounts, since the shares underlying merger and CEF arbitrages were more liquid than convertible bonds or CDS, making them prime "sell" candidates for funds that were desperate to raise capital. We might expect a similar effect in the market for S&P 500 futures, the asset underlying any convergence trade involving options on the S&P 500 futures.

expiration). We construct an index measuring the absolute deviations of domestic equity CEFs from their NAVs. We use absolute deviations instead of discounts because CEFs that trade at premiums are similarly mispriced, and we don't want the negative "discounts" on premium funds to net out the discounts on discounted CEFs. We estimate regressions and assess the ability of our equity CEF mispricing index and the systematic liquidity measure of HPW to improve forecasts of future realized volatility.

We find that lagged values of implied volatility (either Black-Scholes or model-free measures) are statistically significant in forecasts of future volatility, with the Black-Scholes measure performing slightly better than the model-free implied volatility. We find that lagged measures of market-wide liquidity and closed-end fund absolute mispricing are also statistically significant in forecasts of future volatility. A 1% increase in the HPW market-wide liquidity index (refecting worsening liquidity) predicts a 3.51% increase in realized volatility and a 1% increase in the closed-end fund absolute mispricing index forecasts a 4.14% increase in realized volatility.

We show that implied volatility subsumes the HPW index of market-wide liquidity, but not the closed-end fund mispricing index. Using the closed-end fund mispricing index and its lag, along with lagged measures of implied volatility, we are able to explain about 71% of the total variation in realized volatility, significantly better than a naive forecast based only on lagged realized volatility, which produces an $R^2$ of 53%. We conclude that the closed-end fund mispricing index contains additional information about the limits to arbitrage beyond that captured by the index of market-wide liquidity.

The rest of the paper is organized as follows. In Section 2, we summarize the relevant literature on volatility forecasting and limits to arbitrage and motivate the variable choices for our volatility forecasting model. In Section 3 we describe our methodology, especially our estimation procedure for model-free implied volatilities. We present our empirical results in Section 4. Section 5 concludes.

## 2. Literature Review

### *2.1 Volatility Forecasting Models*

Volatility forecasting is the subject of a plethora of academic papers and research reports by practitioners.[3] These papers tend to fall into two general categories: ARCH and GARCH-type models based on past time-series behavior of the realized volatility process, and market variable-based models that use contemporaneous market-determined variables (e.g., option implied volatility) as forward-looking measures of investor expectations. The former attempt to describe a stochastic process that is consistent with past observations on realized volatility, focusing on time-series econometrics rather than the economic fundamentals that underlie observed volatility series, while the latter exploit the fact that expectations of future volatility are a crucial input into option pricing models and are naturally forward-looking. It is generally accepted that forecasts based on measures of implied volatility are superior to ARCH/GARCH-type forecasts, though there is little consensus on how best to compute the implied volatility.

[3] See Poon and Granger, 2003, and references cited therein, as well as more recent papers such as Jiang and Tian, 2005a, b and Anderson and Bondarenko, 2007.

Implied volatilities based on Black-Scholes (1973) dominate the earliest market variable-based volatility forecasting models.[4] These papers generally find that the Black–Scholes implied volatility (BSIV) is superior to historical volatility as a predictor of future realized volatility, but it is biased in that the BSIV tends to exceed future realized volatility. A simple model clarifies this issue.

$$RVol_t = b_0 + b_1 IVol_{t-1} + u_t \qquad (1)$$

Here RVol and IVol are the realized and implied volatilities, respectively, and u is a random error term. If IVol is an unbiased forecast of RVol, then the coefficient, $b_1$, should be statistically equal to 1 and the intercept term, $b_0$, should be equal to zero. In general, early studies produce coefficient estimates of $b_1$ of 0.7 to 0.8, statistically well below 1. Researchers speculated that the bias is due to model mis-specification, microstructure effects, non-synchronous trading, and the existence of the wildcard option (see Figlewski, 1997, among others). Based on the $R^2$s reported in prior studies, we know that IVol explains roughly half of the variation in RVol.

Many prior studies relied on implied volatility measures from options on the OEX (S&P 100) because this was the most liquid index in the 1970s and 1980s. These models suffer from institutional issues that constrain the arbitrage mechanism, allowing option prices to become disconnected from fundamentals. First, it is difficult to trade the underlying 100 stocks in unison as a basket, leaving open a significant possibility of arbitrage persistence. Second, some stocks in the OEX trade rather infrequently implying that prices may be stale. Third, the OEX has a built in wild card option because the market for the underlying stocks closes at 4:00, while the OEX options pits are open until

[4] See, for example, Canina and Figlewski (1993), Lamoureux and Lastrapes (1993), Jorion (1995) and Christensen and Prabhala (1998), among others.

4:15. Finally, a researcher must estimate the expected dividend stream on the underlying basket of stocks over the option's life.

By using S&P 500 futures and futures options, we circumvent these problems. First, the S&P 500 futures and futures options trade in pits that are physically located next to one another, thereby facilitating arbitrage trades. Second, both the options and the underlying futures contracts trade on markets that are open until 4:15 (and all our trades are time-stamped to the second), thereby reducing the problem of non-synchronous pricing and eliminating the wildcard option. Third, the S&P 500 futures contracts are among the most liquid equity futures contracts in the world. Finally, we have trade-by-trade data on the underlying contracts. Thus we are able to compute realized volatility measures based on prices sampled every five minutes. These five-minute returns are serially uncorrelated, which is not the case with the underlying SPX (the index underlying the futures contracts).[5]

More recent volatility forecasting models incorporate computational innovations for realized and implied volatility measures. Andersen et al. (2003) and Anderson, Bollerslev, and Meddahi (2005) show the superiority of using high-frequency data to compute realized volatility, concluding that 5-minute pricing provides a much better estimate of realized volatility than daily pricing. Britten-Jones and Neuberger (2000) derive a *model-free implied volatility* (MFIV) and Jiang and Tian (2005a, b) show that MFIV remains valid in the presence of jumps. They also describe how MFIV can be estimated consistently. We follow Andersen et al. (2003) and Anderson, Bollerslev, and Meddahi (2005) and compute realized volatilities using 5-minute pricing, and we

[5] See Jiang and Tian (2005a).

implement the MFIV algorithm of Jiang and Tian (2005a) to construct our own model-free implied volatilities.

When the Chicago Board Options Exchange (CBOE) introduced their volatility index (VIX) in 1993, they measured volatility according to Black and Scholes (1973); in 2003, they switched to MFIV. When one tests whether implied volatility is an efficient predictor of future realized volatility, one simultaneously tests whether the underlying options pricing model is valid and if option prices are efficient. Unlike BSIV, MFIV does not suffer from this "joint test" problem. However, as a forecasting tool, MFIV is not without its shortcomings.

The Britten-Jones and Neuberger (2000) formula for MFIV takes the form of an integral (sum) of expected square returns over a range of strike prices. Of critical importance is that this expectation is evaluated using risk-neutral probabilities, rather than objective probabilities. If volatility is stochastic and there is a risk premium associated with volatility risk, then the risk-neutral and objective expectations of future squared returns will differ significantly, leading to errors in forecasts of future realized volatility based on MFIV.

### *2.2 Limits to Arbitrage*

If arbitrageurs face constraints on their trading activities, option prices will adjust to reflect these constraints, thereby affecting the risk neutral probability measure and as a consequence, the implied volatility. Limits to arbitrage can be categorized into two types: costs resulting from systematic liquidity constraints or liquidity shocks and costs associated with idiosyncratic illiquidity or event risks specific to a particular asset (such as changes in margin requirements, or restrictions on shorting, costs related to basis risk, uncertainty about portfolio holdings, or unhedgeable risks such as higher moments of risk due to market incompleteness.

In the presence of increased arbitrage costs or risks, arbitrageurs will demand a greater reward in the form of a larger price disparity as compensation for bearing the additional costs/risks. Shleifer and Vishny (1997) argue that it is precisely when deviations from fundamental value are greatest that arbitrageurs become extremely reluctant to execute or maintain arbitrage trades, due to their relatively shallow pockets and short time horizons.

Increased uncertainty or concerns about a jump in volatility increase the costs of arbitrage. Arbitrageurs need to make trades, so their ability to enter and exit trades is paramount. Deuskar (2006) shows that volatility, expected volatility, and illiquidity are very much interrelated and self-reinforcing. Liquidity often dries up when investors expect volatility to increase. Confronted with a mispriced asset that is moving in the wrong direction, an arbitrageur may be forced to liquidate his positions at the worst possible time, a risk Shleifer and Vishny (1997) label as performance-based arbitrage.

We acknowledge that limits to arbitrage affect security prices and result in deviations from fundamental value. We search for proxies that capture the systematic and idiosyncratic risks that arbitrageurs face. We believe that by incorporating these variables in our model of volatility, we can improve the forecasts of future realized volatility.

### *2.3 Funding Liquidity*

HPW argue that the abundance of arbitrage capital during normal times helps smooth out the Treasury yield curve and keep the average yield dispersion low. When yields fall out of line, hedge funds and proprietary traders at investment banks step in to execute relative value and arbitrage trades across various habitats of the Treasury yield curve. The simplicity and transparency of these trades make them very appealing and relatively easy to execute.[6]

During liquidity crises, however, the lack of arbitrage capital forces traders to limit or even abandon their relative value trades, allowing yields to move more freely and resulting in more yield curve noise. HPW argue that this abnormal noise in Treasury yields is a symptom of a market in severe shortage of arbitrage capital. Moreover, if active traders allocate capital across various asset classes, a shortage of arbitrage capital in the Treasury market can quickly spread to other markets. For this reason, we argue that the HPW noise measure proxies for systematic liquidity.

### *2.3 Non-systematic Limits to Arbitrage*

While open-end funds are required to redeem or issue new shares at the reported net asset value (NAV) at the end of each trading day, closed-end funds (CEFs) face no such redemption requirements. Instead, closed-end fund shares trade on equity markets just like stock.[7] Therefore, for closed-end funds we observe both an NAV and a price, and these two quantities are usually different.

---

[6] The arbitrage strategy is to buy Treasuries which yield more than their analog zero coupon yields and sell Treasuries which yield less than their zero coupon counterparts.
[7] Compared to open-end funds, closed-end funds also make greater use of leverage.

For the vast majority of closed-end funds, the share price is typically well below the NAV, resulting in the so-called closed-end fund discount. Many academics view the closed-end fund discount as compelling evidence of investor irrationality, where retail investors (noise traders) drive prices away from fundamental value (NAV).[8] However, numerous rational explanations for the closed-end fund discount exist, including liquidity differences between CEF holdings and CEF shares, CEF distribution policies, CEF portfolio manager skills and compensation, unrealized capital gains, and agency problems.[9] In this paper, we ignore the root causes of discounts (or premiums) and instead focus on what impediments might arise to limit CEF arbitrage trades.

Quite a few hedge funds and mutual funds are engaged in CEF arbitrage. There are also activist hedge funds that target CEFs trading at discounts, hoping to take control of a fund and either liquidate the assets or open-end the fund. These traders represent a powerful force against CEF mispricing. However, as Shleifer and Vishny (1997) and Pontiff (1996) argue, CEF arbitrage does not conform to the academic ideal of costless, riskless arbitrage that requires no capital. These papers argue that costs inhibit CEF arbitrage and deviations from fundamental value are an increasing function of these costs.

Trading costs are not negligible, especially for short positions in the CEF shares or the CEF assets, which are often illiquid. Second, basis risk can be quite substantial because CEF holdings are observed with a lag and only once per quarter (at most). Third,

---

[8] Examples of these sentiment-based explanations of the CEF discount include De Long et al. (1990), Lee, Shleifer and Thaler (1991), and Shleifer and Vishny (1997).

[9] Cherkes, Sagi and Stanton (2009) focus on a liquidity differential; Cherkes, Sagi, and Wang (2009) look at CEF distribution policy; Berk and Stanton (2007) model CEF manager ability and compensation; unrealized gains and agency problems are the focus of Malkiel (1997) and Brennan and Jain (2008), among others.

an increase in volatility will exacerbate the mismatch between an arbitrageur's portfolio and the CEF.

We believe mispricing in the CEF market provides a window into the non-systematic risks that arbitrageurs face. For this reason, we construct an index that measures the absolute deviation of closed-end fund prices from NAV for a subset of domestic, equity-based closed end funds.

## 3. Data and Methodology

### *3.1 Systematic Risks for Arbitrageurs: Funding Liquidity*

Our measure of systematic (funding) liquidity is based on the noise measure of HPW.[10] The authors first estimate a smooth zero-coupon yield curve using daily Treasury security price data. Each Treasury security is then benchmarked to a similar-maturity zero coupon yield. The liquidity index squares and aggregates yield differentials.

### *3.2 Non-systematic Risks for Arbitrageurs: Closed-end Fund Mispricing*

Using the closed-end fund data from Morningstar, we construct our CEF mispricing index as follows. The mispricing on a closed-end fund equals the absolute value of the difference between the closed-end fund price and its net asset value, expressed as a percentage of the net asset value.

$$\text{CEF Mispricing} = |\text{Price} - \text{NAV}| / \text{NAV} \tag{2}$$

On each trading day we compute the CEF mispricing measure for each domestic equity CEF. Our mispricing index is set equal to the arithmetic average of the CEF mispricing measure for all funds (an equally-weighted index) or a weighted average of the CEF

[10] We thank Jun Pan for making these data available on her website. The data span the period 1987 - 2012.

mispricing measure for all funds. Here a fund's weight equals the product of the fund's NAV and its shares outstanding at the end of the previous month.[11]

### *3.3 Realized Volatilities*

We follow Andersen et al. (2003) and Anderson, Bollerslev, and Meddahi (2005) in calculating realized volatility based on returns over 5 minute intervals (an intra-day measure).[12]

### *3.4 Black-Scholes Implied Volatilities*

We collect intraday data on S&P 500 futures contracts traded on the Chicago Mercantile Exchange (CME) during the period 1997 - 2008. We also collect daily closing prices on S&P 500 futures options (also traded at the CME). We measure implied volatility two different ways: the Black-Scholes implied volatility (BSIV) using Whaley's (1986) adjustment for futures; and the model-free implied volatility (MFIV) based on Britten-Jones and Neuberger (2000), and Jiang and Tian (2005a, b).

### *3.5 Estimation of Model-Free Implied Volatilities*

We compute model-free implied volatilities based on Proposition 1 in Jiang and Tian (2005a) which states that the integrated variance (square of volatility) from time 0 to date T is specified by the set of all *call* options expiring at T through the following integral:

$$E\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^\infty \frac{C^F(T,K)-\max(0,F_0-K)}{K^2}dK \qquad (3)$$

where K represents the strike price and $F_t$ represents the forward probability measure. This expression comes from Britten-Jones and Neuberger (2000) and Jiang and Tian

[11] We use monthly data on CEF outstanding shares.
[12] See also Anderson and Bondarenko (2007) and Poteshman (2001).

(2005a) derive this expression under more general assumptions (when asset prices contain jumps). Notice that the numerator of the integrand represents the time value of a call option (call price minus intrinsic value). The integral is taken over a continuum of strikes between 0 and infinity.

The problem with trying to estimate (3) using traded options is that there are only a limited number of contracts being traded at any given time that expire at time T. Thus, in general, one needs to solve something akin to (4) below:

$$E\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_{K_{Min}}^{K_{Max}} \frac{C^F(T,K) - \max(0, F_0 - K)}{K^2} dK \tag{4}$$

where $K_{Min}$ is the lowest traded strike and $K_{Max}$ is the highest traded strike. This is how the VIX index has been estimated since 2003. Another problem arises because there do not exist a continuum of strike prices. Jiang and Tian (2005a) propose two alternatives to deal with the fact that $K_{Min} > 0$ and $K_{Max} < \infty$: truncation and extrapolation. Truncation implies that the information in calls either with strikes less than $K_{Min}$ or greater than $K_{Max}$ are ignored (as the CBOE does with the calculation of the VIX); extrapolation includes at least some of these options in the computation.

Jiang and Tian (2005a) find that the extrapolation method is typically an order of magnitude more accurate than the truncation method. Setting $K_{Min}$ and $K_{Max}$ at least 2 standard deviations away from at-the-money is usually sufficient to ensure accurate results. Both of these conclusions are arrived at via simulations based on estimating the MF volatility for options whose prices stem from a model (Heston model with specific parameter choices). Nonetheless, we set $K_{Min}$ and $K_{Max}$ 4 standard deviations away from the then current level of the underlying futures contract. We also consider the truncation approach in separate tests; our results are qualitatively similar.

Equation (3) needs to be solved numerically. The first step is to deal with the fact that there do not exist call options with a continuum of strikes expiring at T. Once this problem has been dealt with, we focus on the issue of truncation error that stems from the fact that there do not exist options with strikes that surpass $K_{Max,}$ nor options with strikes less than $K_{Min}$. The first problem is one of interpolation, the second, one of extrapolation. We consider extrapolated as well as truncated (i.e., un-extrapolated) solutions.

*3.51 Interpolation Using Cubic Splines*

To solve the problem of lack of a continuum of strikes, we again follow Jiang and Tian (2005a) and use cubic splines. We create a smooth curve that is fitted exactly, based on observed prices, and interpolate using cubic polynomials for values between these observed prices. Jiang and Tian (2005a) argue that given that option values are highly non-linear functions of strike prices, there is precedent in the literature (see Ait-Sahalia and Lo, 1998) to use a curve-fitting algorithm to form a volatility surface and transform it into a price surface, rather than forming a price surface directly.

Each observed price is turned into an implied volatility using the Black-Scholes model. Cubic splines are then applied to form a volatility surface. Then any point on that volatility surface can be converted back to a price using the Black-Scholes model. In this way one can create as fine a grid of option prices as is desired (we use grid increments of one index point, which is more than sufficient to insure accuracy). Note that the Black-Scholes model is only being used as a means to transform prices into volatilities and then back into prices. As such it does not impose an assumption that option prices behave as in the Black-Scholes model.

The spline is a curve-fitting algorithm that imposes some smoothness conditions that enable a researcher to fit a 3rd degree polynomial to a given set of data, along with conditions on the first and second derivatives of the spline to ensure smoothness. Take adjoining intervals at time t and t+1. Since each consecutive interval shares a common point with its predecessor, smoothness necessitates that the spline function itself, its first derivative, and its second derivative from interval t, evaluated at interval t's right endpoint, must equal the comparable terms for the spline in interval t+1, evaluated at its left endpoint to ensure smoothness. Imposing these smoothness conditions solves for the coefficients on the splines in each interval.

*3.52 Extrapolation Beyond Traded Strikes*

In order to extrapolate to strikes that lie outside the range of options traded at any point in time, we again follow Jiang and Tian (2005a) and use the following algorithm. We measure the implied volatility for the option with the lowest strike. We then assume that all *lower* strikes have the same implied volatility as the traded option with the lowest strike and use the Black-Scholes model to estimate call prices for those strikes that fall below the traded range. An analogous procedure can be applied to strikes beyond the highest traded strike.

*3.53 Numerical Integration Procedure*

Armed with a complete set of call prices (for as wide a range of strikes as is deemed necessary), we can go about estimating the integral in (4), which is an approximation to the integral depicted in (3). We use the Trapezoidal Rule to evaluate the integral in (4) numerically, which is straightforward. The "range" of the integral (from the bottom endpoint to the top endpoint) is divided into intervals of 1 index point. For any interval, say between $X_0$ and $X_0 + \alpha$, linearly interpolate between $f(X_0)$ and $f(X_0 + \alpha)$. The four points, $(X_0, 0)$, $(X_0 + \alpha, 0)$, $(X_0 + \alpha, f(X_0 + \alpha))$, and $(X_0, f(X_0))$ form a trapezoid whose area is given by: $\alpha \times \left[ f(X_0) + f(X_0 + \alpha) \right] / 2$. By summing the areas of these trapezoids we get an estimate of the integral. As long as the integrand is fairly smooth and the grid is fine enough then we can make this as accurate as necessary.

Following Jiang and Tian (2005a), we ignore all in-the-money calls, defined as options with strike prices less than 97% of the futures price, (i.e., cases where the given option is more than 3% in the money), and instead focus on out-of-the-money put options (and out-of-the-money call options) to derive our volatility surface. It is well established that in-the-money options are not very liquid and therefore pose considerable hurdles. The out-of-the-money options are far more liquid.

*3.6 Dataset Construction*

We construct our non-telescoping dataset in the following way: we sample option prices at one-month intervals, strategically choosing the interval so as to minimize microstructure effects while preserving consistency of the interval length. We select the Tuesday following the expiration of the previous contract (-we use Wednesday if Tuesday is a holiday), and we focus on the shortest duration contracts. In this way the options contracts we consider all have approximately 20 trading days until expiration. Moreover, we compute future realized volatility over exactly the same interval of approximately 20 trading days. Therefore, there is an exact match between the interval used to estimate future realized volatility and the interval corresponding to the option's remaining life, that we use to impute an implied volatility.[13]

We illustrate our approach using the May 2013 options. These contracts expired on Saturday, May 18, 2013, though the last trading day was Friday, May 17. We estimate implied volatility based on closing prices as of Tuesday, May 21, 2013. On that date, the shortest maturity contracts expire on Saturday, June 22, 2013 (24 trading days hence). $RVOL_t$, is our estimate of realized volatility over the interval 5/21/2013 – 6/22/2013 and $IMPVOL_{t-1}$ is our estimate of implied volatility based on option prices as of 5/21/2013. We annualize all volatility estimates to mitigate the effects of slight variations in the interval lengths. The HPW liquidity index and our CEF mispricing index are sampled on the same date as implied volatility.

Finally, as it turns out, all the series we consider in our volatility forecasting model have a fair amount of serial correlation, particularly our limits to arbitrage proxies. Thus, inclusion of lags and an examination of first differences are critical steps in our analysis to insure that our regressions are well-specified.

Armed with our measures of systematic and non-systematic risks faced by arbitrageurs, we estimate the following equation.

$$RVol_t = b_0 + b_1 RVol_{t-1} + b_2 IVol_{t-1} + b_3 LIQUID_{t-1} + b_4\, CEFMIS_{t-1} + u_t \qquad (5)$$

Here, CEFMIS is one of our closed-end fund mispricing indexes and LIQUID is the liquidity/noise measure of HPW. If our hypothesis is correct and our measures of

[13] In contrast, the CBOE's volatility index (a.k.a. the "VIX") is constructed to forecast volatility over a fixed interval of 30 days. To arrive at this 30 day maturity, the CBOE averages implied volatilities from two adjacent contracts, one expiring in less than 30 days and one expiring in more than 30 days. Additionally, during the expiration week of the nearest term contract, the CBOE ignores the shortest maturity contract and instead averages long and short positions in the second and third shortest maturity contracts, such that the weighted average maturity remains 30 days. This approach reflects fairly strict assumptions about the time-varying properties of volatility.

systematic and idiosyncratic risks are capturing risks faced by arbitrageurs, then we should observe one or more of the following outcomes. The coefficient estimate for $b_2$ should approach 1, the coefficient estimates for $b_3$ and/or $b_4$ should be significant, and our $R^2$ values should increase.

## 4. Regression Estimates

We present summary statistics for our data in Table 1. Over the sample period 1997 – 2008, realized volatility (RVOL) ranges from a low of 7.06% to a high of 72.90%, with a mean value of 18.36%. In contrast, both measures of implied volatility (IMPVOL) traverse narrower ranges, but exhibit higher mean values. This result comports with our hypothesis that option prices contain a premium that reflects limits to arbitrage.

The market-wide liquidity measure (LIQUID) ranges from a low of 1.05 basis points (bp) to a high of 17.03 bp, with a mean value of 2.38 bp. Here, higher values denote reduced liquidity. Rescaling these values in terms of deviations from the mean, we obtain a range of $\mu - 0.92\sigma$ to $\mu + 5.79\sigma$, similar to the realized volatility series which travels between $\mu - 1.13\sigma$ to $\mu + 5.43\sigma$. In contrast, our CEF mispricing indices travel a narrower range (e.g., $\mu - 1.79\sigma$ to $\mu + 2.66\sigma$, in the case of our value-weighted index).

Table 2 summarizes the correlation structure of our data. We examine levels data in Panel A and first-differences in Panel B since all of our variables are AR(1). For the levels data, all pairwise correlations are statistically significant at the 1% level, with correlation values ranging from 39% to 95%. BSIV and MFIV are highly correlated and both measures of implied volatility independently have about an 80% correlation with realized volatility. Our measures of market-wide liquidity and unhedgeable/other risks captured by closed-end fund price deviations from NAV (CEFMISP-VW and CEFMISP-EW) display much lower pairwise correlations. Similar results obtain for first differences, shown in Panel B of Table 2. All of the pair-wise correlation values are statistically significant at the 2% level, except for the correlation between realized volatility changes and market-wide liquidity changes.

In Tables 3 – 5, we examine various specifications of a volatility forecasting model that uses the implied volatilities from option prices and measures of market-wide liquidity and other risks to predict future realized volatility. Our basic model is described in equation (6) below:

$$RVOL_t = a + bRVOL_{t-1} + \sum_{i=1}^{2} c_i IMPVOL_{t-i} + d_i LIQUID_{t-i} + e_i ORISKS_{t-i} + \varepsilon_t \quad (6)$$

Here, $RVOL_t$ is the realized volatility of S&P 500 futures returns over the period [t-1, t], based on price data sampled at 5-minute intervals. $IMPVOL_{t-1}$ is the implied volatility of the S&P 500 futures returns, based on daily closing prices for S&P 500 futures options, expiring at time t, and sampled at time t-1 (see Section 3). We compute implied volatility two different ways: the Black-Scholes implied volatility (BSIV) using Whaley's (1986) adjustment for futures; and the model-free implied volatility (MFIV) based on Britten-Jones and Neuberger (2000), and Jiang and Tian (2005a, b). $LIQUID_{t-1}$ is a lagged measure of market-wide liquidity (actually an illiquidity index), based on HPW, and $ORISKS_{t-1}$ is a lagged measure of closed-end fund absolute deviations from net asset value (NAV), either value-weighted (CEFMISP-VW) or equally-weighted (CEFMISP-EW), averaged across the universe of US domestic equity closed-end funds. In our tests of equation (6), we use non-overlapping monthly data for the period 1997 – 2008. We have 144 observations in total.

Table 3 reports coefficient estimates and t-statistics for specifications of equation (6) based on the Model-free implied volatility (MFIV). Panel A includes models with one or two system variables; Panel B includes models with more than two system variables. In Panel A, we see that MFIV does a good job forecasting future realized volatility (Model 2), with an adjusted $R^2$ of 63%. However, the residuals are serially correlated, casting doubt on our t-stat and $R^2$ measures. Adding additional regressors leads to well-behaved error terms and higher explanatory power. Realized volatility is positively related to the first-lag measure of market-wide liquidity (Model 5). A 1% increase in market-wide liquidity forecasts a 3.51% increase in realized volatility. Realized volatility is also positively related to the first-lag measure of closed-end fund mispricing and negatively related to an additional lag of the closed-end fund mispricing (Model 7). A 1% increase in the closed-end fund absolute mispricing index forecasts a 4.14% increase in realized volatility.

In Model 4, we see that lagged measures of implied volatility on the right hand side largely subsume the lagged measure of realized volatility; the coefficient estimate on lagged realized volatility drops from 0.73 to 0.28 and the t-statistic falls from 12.37 to 2.27. In contrast, lagged realized volatility subsumes the liquidity index (compare Models 5 and 6). Interestingly, our closed-end fund mispricing index retains its forecasting power (compare Models 7 and 8). The evidence in Table 3 confirms that our proxies for systematic and idiosyncratic risks are capturing aspects of the limits to arbitrage.

In Panel B, Models 9 - 14 test whether our systematic and idiosyncratic risk proxies have explanatory power beyond the information contained in lagged measures of implied volatility and realized volatility.  Not surprisingly, since the liquidity index (LIQUID) was already shown to be redundant in the presence of lagged historical volatility, the coefficients on LIQUID and its lag are insignificant in models 9 - 12. However, our closed-end fund mispricing index retains its predictive power.  In Models 13 and 14, which combine lagged measures of implied volatility, realized volatility and closed-end fund mispricing, we are able to explain about 71% of the variation in realized volatility, significantly better than a naive forecast based on lagged realized volatility alone (Model 1 in Panel A), which produces an $R^2$ of 53%.

Table 4 reports coefficient estimates and t-statistics for specifications of equation (6) based on the Black-Scholes implied volatility (BSIV).  As in Table 3, we see that BSIV does a good job forecasting future realized volatility (Model 1), with an adjusted $R^2$ of 64%.  However, the residuals are serially correlated, casting doubt on our t-stat and $R^2$ measures.  Adding additional regressors leads to well-behaved error terms and higher explanatory power.  As we saw in Table 3, the closed-end fund mispricing index retains its explanatory power in the presence of the other regressors, while LIQUID is subsumed by lagged historical volatility.  Using all regressors, including the closed-end fund mispricing index (Model 8), we are able to explain about 71% of the variation in realized volatility, significantly better than a naive forecast based on lagged realized volatility only (Model A), which produces an $R^2$ of 53%. Comparing Table 4 to Table 3, we see that BSIV does a slightly better job forecasting realized volatility than does MSIV.

In an effort to induce symmetry in our time-series data, we alter equation (6) by taking logarithmic transformations of the dependent and independent variables. Our results, presented in Table 5, are qualitatively similar, with the adjusted $R^2$ values slightly higher.[14]

We also examine first-differences of our time-series data in Table 6. We estimate the following model:

$$\Delta RVOL_t = b_0 + b_1 \Delta MFIV_{t-1} + b_2 \Delta LIQUID_{t-1} + b_3 \Delta ORISKS_{t-i} + \varepsilon_t \qquad (7)$$

Here $\Delta RVOL_t$ is defined as $RVOL_t$ - $RVOL_{t-1}$, and $\Delta MFIV_{t-1}$ is defined as $MFIV_{t-1}$ - $MFIV_{t-2}$; other variables are similarly defined. We report univariate results in Models 1 - 3. Here we see that changes in implied volatility and changes in the closed-end fund mispricing index forecast changes in realized volatility, with 12% to 14% of the variation in realized volatility changes explained. In contrast, innovations in the liquidity index have no predictive power.

In Model 4, we see that innovations in the liquidity index work in tandem with implied volatility innovations to explain a combined 20% of the variation in realized volatility innovations. Similarly, innovations in the closed-end fund mispricing index also work in tandem with implied volatility innovations to explain 20% of the variation in changes in realized volatility. Finally, Model 7 shows that all three innovation variables (implied volatility, liquidity and closed-end fund mispricing) are significant in explaining innovations in realized volatility; here the adjusted $R^2$ rises to 27%.

We perform several tests to assess the marginal contributions of some of our independent variables when it comes to forecasting realized volatility. This is particularly important for our proxies of systematic and idiosyncratic risks. We start by orthogonalizing the liquidity index with respect to implied volatility and using the

---

[14] Note that the results in Table 5 are based on MFIV, but the results are qualitatively similar if based on BSIV.

residuals from this initial regression in place of the liquidity index in a second model that predicts realized volatility. Our approach is described in Equation (8).

$$\begin{aligned} LIQUID_t &= a_0 + a_1 MFIV_t + \gamma_t \\ RVOL_t &= b_0 + b_1 RVOL_{t-1} + b_2 MFIV_{t-1} + b_3 MFIV_{t-2} + b_4 \gamma_t + b_5 \gamma_{t-1} + \varepsilon_t \end{aligned} \tag{8}$$

We then repeat this procedure, but in the first step we orthogonalize implied volatility with respect to the liquidity index. With either approach, we test whether the residuals (and lagged residuals) from the first step have explanatory power in the second regression.

Our results appear in Table 7. Panel A reports the results of orthogonalizing the liquidity index on implied volatility; Panel B reports the results of orthogonalizing implied volatility on the liquidity index.

In Panel A, the residuals have no explanatory power, whereas our measures of implied volatility and lagged realized volatility are statistically significant. These results comport with Model 10 from Table 3.

In contrast, the residuals (and the lagged residuals) are statistically significant at the 1% level in Panel B, as is the lagged liquidity index. We can reasonably conclude that MFIV contains additional information about the limits to arbitrage beyond that which is captured by funding constraints, but the liquidity index of HPW contains no volatility-relevant information beyond that already reflected in implied volatility and lagged realized volatility.

We repeat this analysis in Table 8 but this time we expand the list of regressors to include our closed-end fund mispricing index as well as its lag. As we saw in Model 14 from Table 3, we see that the mispricing index has statistically significant power to forecast future realized volatility. Once again, we can reasonably conclude that MFIV contains additional information about the limits to arbitrage beyond that which is captured by funding constraints, but the liquidity index of HPW contains no volatility-relevant information beyond that already reflected in implied volatility and lagged realized volatility. Moreover, our CEF mispricing index contains volatility relevant information beyond that contained in either MFIV or the liquidity index of HPW.

Figure 1 illustrates these relationships. First, any volatility-relevant information in the liquidity index of HPW is subsumed by MFIV. Second, our CEF mispricing index and MFIV each contains volatility-relevant information that is unique, distinct from each other, and distinct from the information in the HPW liquidity index. These findings underscore our conceptualization of the systematic and idiosyncratic effects of limits to arbitrage. In so far as the liquidity measure of HPW is systematic, innovations in that index reflect information that is relevant to all arbitrageurs and are reflected to some extent in all assets. However, in addition to these systematic effects, our CEF mispricing index reflects conditions specific to those assets that may or may not be relevant in other markets. The same can be said of MFIV.

Finally, in results not reported in tabular form, we also regress our closed-end fund mispricing index (CEFMISP-VW) on our market-wide measure of liquidity (LIQUID). We test whether the residuals from this regression are statistically significant predictors of realized volatility. It turns out they are. When we reverse the procedure, regressing the market-wide measure of liquidity (LIQUID) on the closed-end fund mispricing index (CEFMISP-VW), we find that the residuals from this regression are not statistically significant predictors of realized volatility. We conclude that CEFMISP-VW contains additional information about the limits to arbitrage beyond those captured by funding constraints.

These results are robust to different measures of implied volatility, such as BSIV, as well as models using logarithmic transformations. Moreover, all results are qualitatively similar if we replace our value-weighted mispricing index with the equally-weighted mispricing index.

## 5. Conclusions

In this paper, we develop a model to forecast the volatility of S&P 500 futures returns. We examine several potential predictors, including realized volatility, implied volatility, a measure of market-wide liquidity and a measure of closed-end fund mispricing, arguing that the last two variables capture the effects of systematic and idiosyncratic limits to arbitrage, respectively. We estimate realized volatilities based on S&P 500 futures price data sampled at 5-minute intervals. We use end-of-day S&P 500 futures options prices to compute both Black-Scholes implied volatilities, based on Whaley (1986), and model-free implied volatilities, based on Britten-Jones and Neuberger (2000). Our measure of market-wide liquidity comes from Hu, Pan, and Wang (2012), whose liquidity (noise) measure is computed as the mean squared deviation of Treasury securities from the zero coupon yield curve. The market-wide liquidity index captures funding constraints that arbitrageurs face. It is one of the limits to arbitrage.

We conjecture that there are additional limits to arbitrage that are better classified as idiosyncratic. Arbitrageurs face costs associated with event risks specific to a particular asset (such as changes in margin requirements, or restrictions on shorting). They also face costs related to basis risk, uncertainty about portfolio holdings, or unhedgeable risks (such as higher moments of risk) due to market incompleteness. We develop an index that measures the absolute price deviations of equity closed-end funds from their net asset values. We argue that this index proxies for some of these additional limits to arbitrage.

We find that lagged values of implied volatility are statistically significant in forecasts of future volatility. We find that lagged measures of market-wide liquidity and closed-end fund absolute mispricing are also statistically significant in forecasts of future volatility.

We show that our closed-end fund mispricing index contains important information that forecasts future realized volatility, while our systematic measure of liquidity, the liquidity index of Hu et al. (2012), contains no information beyond that already contained in implied volatility. Using the closed-end fund mispricing index, along with lagged measures of implied volatility and realized volatility, we are able to explain about 71% of the total variation in realized volatility, significantly better than a naive forecast based only on lagged realized volatility, which produces an $R^2$ of 53%. We conclude that the closed-end fund mispricing index contains additional information about the limits to arbitrage beyond that captured by the index of market-wide liquidity.

## References


Adrian, T. and H. Shin (2010). Liquidity and Leverage. *Journal of Financial Intermediation* 19, 418-437.

Ait-Sahalia, Y. and A. Lo, 1998. Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance* 53, 499-547.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys, 2003. Modeling and Forecasting Realized Volatility. *Econometrica* 71, 579-625.

Anderson, T. and O. Bondarenko, 2007. Construction and Interpretation of Model-Free Implied Volatility. CREATES Research Paper #2007-24.

Bakshi, G. and N. Kapadia, 2003. Delta-Hedged Gaines and the Negative Volatility Risk Premium. *Review of Financial Studies* 16, 527-566.

Ben-David, I., F. Franzoni, and R. Moussawi, 2012. Hedge Fund Stock Trading in the Financial Crisis of 2007-2009. *Review of Financial Studies*, forthcoming.

Berk, J., and R. Stanton, 2007. Managerial Ability, Compensation, and the Closed-End Fund Discount. *Journal of Finance* 57, 529-555.

Black, F. and M. Scholes, 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637-654.

Bollerslev, T., M. Gibson, and H. Zhou, 2006. Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option Implied and Realized Volatilities. Duke University working paper.

Britten-Jones, M. and A. Neuberger, 2000. Option Prices, Implied Price Processes, and Stochastic Volatility. *Journal of Finance* 55, 839-866.

Canina, L., and S. Figlewski. The Informational Content of Implied Volatility. *Review of Financial Studies*, 6 (1993), 659–681.

Carr, P. and L. Wu, 2009. Variance Risk Premiums. *Review of Financial Studies* 22, 1311-1343.

Cherkes, M., J. Sagi, and R. Stanton, 2008. A Liquidity-Based Theory of Closed-End Funds. *Review of Financial Studies* 22, 257-297.

Cherkes, M., J. Sagi, and J. Wang. A Neoclassical Model of Managed Distribution Plans: Theory and Evidence. Columbia University working paper.

Chernov, M., 2002. Implied Volatilities as Forecasts of Future Volatility, Time-Varying Risk Premia, and Return Variability. Columbia University working paper.

Christensen, B. J., and N. R. Prabhala, 1998. The Relation Between Implied and Realized Volatility. *Journal of Financial Economics* 50, 125-150.

DeLong, B., A. Shleifer, L.Summers, and R. Waldmann, 1990. Noise Trader Risk in Financial Markets. *Journal of Political Economy* 97, 703-738.

Deuskar, P., 2006. Extrapolative Expectations: Implications for Volatility and Liquidity. University of Illinois working paper.

Duffie, D., 2010. Presidential address: Asset Price Dynamics with Slow Moving Capital. *Journal of Finance* 65, 1237-1267.

Figelwski, S., 1997. Forecasting Volatility. *Financial Markets, Institutions, and Instruments* 6, NYU Salomon Center, 1-88.

Gorton, G. and A. Metrick, 2012. Securitized Banking and the Run on Repo. *Journal of Financial Economics* 104, 425–451.

Heston, S., 1993. A Closed-Form Solution for Options with Stochastic Volatility withApplications to Bond and Currency Options. *Review of Financial Studies* 6, 327-343.

Hu, X., J. Pan, and J. Wang, 2012. Noise as Information for Illiquidity. *Journal of Finance,* forthcoming.

Hull, J. and A. White, 1987. The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance* 42, 281-300.

Jiang, G. and Y. Tian, 2005a. The Model-Free Implied Volatility and its Information Content. *Review of Financial Studies* 18, 1305-1342.

Jiang, G. and Y. Tian, 2005b. Gauging the ‘Investor Fear Gauge’: Implementation Problems in the CBOE’s New Volatility Index and a Simple Solution. University of Arizona Working Paper.

Jorion, P., 1995. Predicting Volatility in the Foreign Exchange Market. *Journal of Finance* 50, 507-528.

Lamoureux, C. and W. Lastrapes, 1993. Forecasting Stock Return Variance: Toward an Understanding of Stochastic Implied Volatilities. *Review of Financial Studies* 6, 293-326.

Lee, C. M. C., A. Shleifer, and R. H. Thaler, 1991. Investor Sentiment and the Closed-end Fund Puzzle. *Journal of Finance* 46, 76–110.

Nohel, T., Todd, S., and Z. Wang, 2013. Leverage in Portfolio Management. University of Oregon working paper.

Pontiff, J., 1996. Costly Arbitrage: Evidence from Closed-End Funds. *Quarterly Journal of Economics* 111, 1135-1151.

Poon, A. and C. Granger, 2003. Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature* 41, 478-539.

Poteshman, A, 2001. Forecasting Future Volatility from Option Prices. University of Illinois at Urbana-Champaign working paper.

Shleifer, A. and R. Vishny, 1997. The Limits of Arbitrage. *Journal of Finance* 52, 35-55.

Todorov, V., 2009. Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, forthcoming.

Whaley, R., 1986. Valuation of American futures options: Theory and empirical tests. *Journal of Finance* 41, 127-150.

## Table 1

Summary Statistics

RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computed using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, as in Jiang and Tian (2005a), $CEFMIS_t$ is the average of the absolute value of the discount on equity closed-end funds at time t (either value-weighted or equally-weighted), computed as |(Price – NAV)/NAV|, and $LIQ_t$ is the noise measure of Hu et al (2012) based on mean deviations from the zero-coupon yield curve at time t. Δs represent changes in the relevant variable, which account for the reduction in N by 1. Min, Max, Std Dev, Mean, and N are the minimum value, the maximum value, the standard deviation of values from the mean, the mean, and the number of observations, respectively.

| Variable | N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|---|
| RVOL | 144 | 18.36% | 10.14% | 7.06% | 72.90% |
| BSIV | 144 | 19.35% | 8.13% | 8.41% | 58.94% |
| MFIV | 144 | 20.58% | 8.45% | 9.12% | 62.97% |
| LIQUID | 144 | 3.26 bp | 2.38 bp | 1.05 bp | 17.03 bp |
| CEFMIS-VW | 144 | 11.71% | 3.25% | 5.92% | 20.36% |
| CEFMIS-EW | 144 | 11.37% | 3.04% | 6.97% | 19.12% |
| ΔRVOL | 143 | -0.10% | 7.55% | -33.36% | 36.42% |
| ΔBSIV | 143 | 0.22% | 5.09% | -16.48% | 24.86% |
| ΔMFIV | 143 | 0.17% | 5.51% | -19.06% | 27.13% |
| ΔLIQUID | 143 | 0.1 bp | 0.91 bp | -1.69 bp | 6.66 bp |
| ΔCEFMIS-VW | 143 | 0.021% | 1.13% | -3.35% | 5.58% |
| ΔCEFMIS-EW | 143 | 0.026% | 0.89% | -2.65% | 4.56% |

# Table 2
Correlations

## Panel A – Levels

CEFMIS is our CEF absolute deviation index (Equally-weighted (EW) or Value-weighted (VW)), MFIV is our model-free volatility measure, BSIV is the CME's implied volatility measure based on Black-Scholes, LIQUID is the Noise measure of Hu et al. (2012), and RVOL is the realized volatility. P-values are in parentheses underneath each correlation. RVOL is forward-looking over the interval corresponding to the expiration date of the options. * and ** indicate significance at the 5% and 1% level, respectively.

| | CEFMIS-EW | CEFMIS-VW | MFIV | BSIV | LIQUID | RVOL |
|---|---|---|---|---|---|---|
| CEFMIS-EW | 1 | 0.95270** (0.0000) | 0.50764** (0.0000) | 0.53652** (0.0000) | 0.41313** (0.0000) | 0.45744** (0.0000) |
| CEFMIS-VW | 0.95270** (0.0000) | 1 | 0.53905** (0.0000) | 0.56519** (0.0000) | 0.38845** (0.0000) | 0.48263** (0.0000) |
| MFIV | 0.50764** (0.0000) | 0.53905** (0.0000) | 1 | 0.99058** (0.0000) | 0.72080** (0.0000) | 0.79726** (0.0000) |
| BSIV | 0.53652** (0.0000) | 0.56519** (0.0000) | 0.99058** (0.0000) | 1 | 0.72736** (0.0000) | 0.80160** (0.0000) |
| LIQUID | 0.41313** (0.0000) | 0.38845** (0.0000) | 0.72080** (0.0000) | 0.72736** (0.0000) | 1 | 0.59758** (0.0000) |
| RVOL | 0.45744** (0.0000) | 0.48263** (0.0000) | 0.79726** (0.0000) | 0.80160** (0.0000) | 0.59758** (0.0000) | 1 |

## Panel B – Changes

ΔCEFMIS is the change in our CEF absolute deviation index (Equally-weighted (EW) or Value-weighted (VW)), ΔMFIV is the change in our model-free volatility measure, ΔBSIV is the change in the CME's implied volatility measure based on Black-Scholes, ΔLIQUID is the change in the Noise measure of Hu et al. (2012), and ΔRVOL is change in the realized volatility. P-values are in parentheses underneath each correlation. RVOL is forward-looking over the interval corresponding to the expiration date of the options. * and ** indicate significance at the 5% and 1% level, respectively.

| | ΔCEFMIS-EW | ΔCEFMIS-VW | ΔMFIV | ΔBSIV | ΔLIQUID | ΔRVOL |
|---|---|---|---|---|---|---|
| ΔCEFMIS-EW | 1 | 0.87028** (0.0000) | 0.42487** (0.0000) | 0.35022** (0.0000) | 0.27449** (0.0010) | 0.35347** (0.0000) |
| ΔCEFMIS-VW | 0.87028** (0.0000) | 1 | 0.37497** (0.0000) | 0.32110** (0.0001) | 0.20685* (0.0139) | 0.35897** (0.0000) |
| ΔMFIV | 0.42487** (0.0000) | 0.37497** (0.0000) | 1 | 0.96998** (0.0000) | 0.41745** (0.0000) | 0.39194** (0.0000) |
| ΔBSIV | 0.35022** (0.0000) | 0.32110** (0.0001) | 0.96998** (0.0000) | 1 | 0.42033** (0.0000) | 0.38575** (0.0000) |
| ΔLIQUID | 0.27449** (0.0010) | 0.20685* (0.0139) | 0.41745** (0.0000) | 0.42033** (0.0000) | 1 | -0.07039 (0.4172) |
| ΔRVOL | 0.35347** (0.0000) | 0.35897** (0.0000) | 0.39194** (0.0000) | 0.38575** (0.0000) | -0.07039 (0.4172) | 1 |

## Table 3: Panel A

Individual Variable Effects

Estimation of the following regression:

$$RVOL_t = a + b\,RVol_{t-1} + \sum_{i=1}^{2} c_i MFIV_{t-i} + d_i LIQUID_{t-i} + e_i ORISKS_{t-i} + \varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|, and $LIQUID_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| a | 0.0490** (3.96) | -0.0106 (-0.78) | 0.0048 (0.34) | 0.0190 (1.25) | 0.1077** (8.61) | 0.0537** (4.07) | 0.0305 (1.14) | 0.0310 (1.54) |
| b | 0.7277** (12.37) | | | 0.2786* (2.27) | | 0.6402** (7.09) | | 0.6352** (10.18) |
| $c_1$ | | 0.9435** (15.52) | 1.1974** (12.52) | 0.9757** (7.20) | | | | |
| $c_2$ | | | -0.332** (-3.37) | -0.432** (-4.01) | | | | |
| $d_1$ | | | | | 3.5110** (4.51) | 1.0832 (1.43) | | |
| $d_2$ | | | | | -1.223 (-1.36) | -0.758 (-0.97) | | |
| $e_1$ | | | | | | | 4.1448** (6.55) | 3.0001** (5.97) |
| $e_2$ | | | | | | | -2.838** (-4.45) | -2.699** (-5.54) |
| Adj $R^2$ | 0.532 | 0.633 | 0.659 | 0.671 | 0.358 | 0.533 | 0.322 | 0.626 |

## Table 3: Panel B

Variable Interactions

Estimation of the following regression:

$$RVOL_t = a + b\,RVol_{t-1} + \sum_{i=1}^{2} c_i MFIV_{t-i} + d_i LIQUID_{t-i} + e_i ORISKS_{t-i} + \varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|, and $LIQUID_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | Model 9 | Model 10 | Model 11 | Model 12 | Model 13 | Model 14 |
|---|---|---|---|---|---|---|
| a | 0.0210<br>(1.37) | 0.0213<br>(1.38) | -0.002<br>(-0.12) | 0.0090<br>(0.46) | -0.005<br>(-0.23) | 0.0076<br>(0.39) |
| b | 0.2472<br>(1.93) | 0.2767*<br>(2.05) | 0.2498<br>(1.85) | 0.295*<br>(2.28) | 0.2690*<br>(2.21) | 0.2908*<br>(2.50) |
| $c_1$ | 0.9604**<br>(7.03) | 0.9651**<br>(7.03) | 0.9221**<br>(6.67) | 0.7908**<br>(5.76) | 0.9383**<br>(6.90) | 0.7918**<br>(5.84) |
| $c_2$ | -0.442**<br>(-4.08) | -0.489**<br>(-3.84) | -0.504**<br>(-3.98) | -0.387**<br>(-3.08) | -0.461**<br>(-4.27) | -0.347**<br>(-3.23) |
| $d_1$ | 0.2811<br>(0.86) | -0.154<br>(-0.22) | 0.0617<br>(0.09) | -0.166<br>(-0.25) | | |
| $d_2$ | | 0.5448<br>(0.70) | 0.3458<br>(0.45) | 0.3884<br>(0.52) | | |
| $e_1$ | | | 0.3432<br>(1.80) | 1.9471**<br>(4.03) | 0.3366<br>(1.80) | 1.984**<br>(4.17) |
| $e_2$ | | | | -1.696**<br>(3.58) | | -1.730**<br>(-3.73) |
| Adj $R^2$ | 0.670 | 0.669 | 0.675 | 0.702 | 0.676 | 0.706 |

# Table 4

Estimation of the following regression:

$$RVOL_t = a + b\,RVol_{t-1} + \sum_{i=1}^{2} c_i BSIV_{t-i} + d_i LIQUID_{t-i} + e_i ORISKS_{t-i} + \varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, BSIV is the implied volatility computed using the Whaley model for futures options (based on Black-Scholes), $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|, and $LIQUID_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| a | 0.0490** (3.96) | -0.010 (-0.75) | 0.0048 (0.35) | 0.0171 (1.15) | 0.0210 (1.37) | 0.0188 (1.24) | -0.001 (-0.02) | 0.0116 (0.61) |
| b | 0.7277** (12.37) | | | 0.2526* (2.07) | 0.2472 (1.93) | 0.2538 (1.90) | 0.2440 (1.83) | 0.2817* (2.23) |
| $c_1$ | | 1.0001** (15.75) | 1.2688** (12.12) | 1.0624** (7.39) | 0.9604** (7.01) | 1.0562** (7.23) | 1.0057** (6.74) | 0.8818** (6.11) |
| $c_2$ | | | -0.351** (-3.24) | -0.449** (-3.83) | -0.442** (-4.08) | -0.505** (-3.65) | -0.528** (-3.81) | -0.403** (-3.00) |
| $d_1$ | | | | | 0.2811 (0.86) | -0.177 (-0.25) | -0.005 (-0.01) | -0.290 (-0.43) |
| $d_2$ | | | | | | 0.5052 (0.65) | 0.3619 (0.47) | 0.4472 (0.61) |
| $e_1$ | | | | | | | 0.2946 (1.51) | 2.0599** (4.38) |
| $e_2$ | | | | | | | | -1.873** (-4.08) |
| Adj $R^2$ | 0.532 | 0.640 | 0.664 | 0.672 | 0.670 | 0.669 | 0.673 | 0.708 |

# Table 5

Estimation of the following regression:

$$Log(1+RVOL_t)=a+b\,Log(1+RVol_{t-1})+\sum_{i=1}^{2}c_i Log(1+MFIV_{t-i})+d_i Log(1+LIQUID_{t-i})+e_i Log(1+ORISKS_{t-i})+\varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|, and $LIQUID_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| a | 0.0429** (4.03) | -0.011 (-0.95) | 0.0006 (0.05) | 0.0119 (0.91) | 0.0137 (1.037) | 0.0140 (1.06) | -0.004 (-0.02) | 0.0041 (0.25) |
| B | 0.7354** (12.68) | | | 0.2556* (2.13) | 0.2309 (1.85) | 0.2533 (1.95) | 0.2266 (1.74) | 0.2597* (2.08) |
| $c_1$ | | 0.9537** (16.18) | 1.1802** (12.29) | 0.9843** (7.48) | 0.9706** (7.29) | 0.975** (7.30) | 0.9327** (6.93) | 0.8191** (6.15) |
| $c_2$ | | | -0.292** (-2.95) | -0.389** (-3.57) | -0.396** (-3.62) | -0.433** (-3.45) | -0.449** (-3.59) | -0.334** (-2.69) |
| $d_1$ | | | | | 0.1919 (0.74) | -0.108 (-0.19) | 0.0534 (0.10) | -0.156 (-0.29) |
| $d_2$ | | | | | | 0.3757 (0.61) | 0.2274 (0.37) | 0.2943 (0.50) |
| $e_1$ | | | | | | | 0.2966 (1.78) | 1.6477** (3.92) |
| $e_2$ | | | | | | | | -1.433** (-3.48) |
| Adj $R^2$ | 0.544 | 0.652 | 0.671 | 0.681 | 0.680 | 0.679 | 0.684 | 0.709 |

## Table 6

Forecasting Volatility Changes

Estimation of the following regression:

$$\Delta RVOL_t = b_0 + b_1 \Delta MFIV_{t-1} + b_2 \Delta LIQUID_{t-1} + b_3 \Delta ORISKS_{t-1} + \varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|, and $LIQUID_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|
| $b_0$ | -0.002 (-0.38) | -0.000 (-0.05) | -0.001 (-0.19) | 0.0002 (0.03) | -0.002 (-0.36) | 0.003 (0.06) | 0.0006 (1.14) |
| $b_1$ | 0.5715** (4.82) | | | 0.7474** (5.93) | 0.4468** (3.69) | | 0.6269** (5.03) |
| $b_2$ | | -0.583 (-0.81) | | -2.339** (-3.32) | | -1.271 (-1.87) | -2.564** (-3.80) |
| $b_3$ | | | 2.3900** (4.44) | | 1.7453** (3.21) | 2.6070** (4.77) | 1.923** (3.70) |
| Adj $R^2$ | 0.142 | -0.003 | 0.122 | 0.203 | 0.199 | 0.139 | 0.272 |

# Table 7

## Panel A

Estimation of the following regression:

$$LIQUID_t = a_0 + a_1\, MFIV_t + \gamma_t$$
$$RVOL_t = b_0 + b_1 RVOL_{t-1} + b_2 MFIV_{t-1} + b_3 MFIV_{t-2} + b_4\gamma_t + b_5\gamma_{t-1} + \varepsilon_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, and $LIQ_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. Note that this is a 2-stage procedure: in the first stage LIQUID is regressed on MFIV, and the residuals from that regression, $\gamma_t$, are used as regressors in place of LIQUID so that the regressors are orthogonal by construction. We also include a lag of $\gamma_t$, as a regressor. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | $b_0$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|
| Model 1 | 0.0178 (1.17) | 0.2767* (2.05) | 0.9340** (5.01) | -0.379** (-2.69) | -0.154 (-0.22) | 0.5448 (0.70) | 0.669 |

## Panel B

Estimation of the following regression:

$$MFIV_t = a_0 + a_1\, LIQUID_t + \nu_t$$
$$RVOL_t = b_0 + b_1 RVOL_{t-1} + b_2\nu_t + b_3\nu_{t-1} + b_4 LIQUID_{t-1} + b_5 LIQUID_{t-2} + u_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, and $LIQ_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. Note that this is a 2-stage procedure: in the first stage LIQUID is regressed on MFIV, and the residuals from that regression, $\nu_t$, are used as regressors in place of LIQUID so that the regressors are orthogonal by construction. We also include a lag of $\nu_t$, as a regressor. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | $b_0$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|
| Model 2 | 0.0793** (5.33) | 0.2767* (2.05) | 0.9651** (7.03) | -0.489** (-3.84) | 2.3292** (3.10) | -0.714 (-1.07) | 0.669 |

## Table 8

### Panel A

Estimation of the following regression:

$$MFIV_t = a_0 + a_1\, LIQUID_t + \nu_t$$
$$RVOL_t = b_0 + b_1 RVOL_{t-1} + b_2 MFIV_{t-1} + b_3 MFIV_{t-2} + b_4\gamma_t + b_5\gamma_{t-1} + b_6 ORISKS_{t-1} + b_7 ORISKS_{t-2} + u_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, and $LIQ_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. Note that this is a 2-stage procedure: in the first stage LIQUID is regressed on MFIV, and the residuals from that regression, $\gamma_t$, are used as regressors in place of LIQUID so that the regressors are orthogonal by construction. We also include a lag of $\gamma_t$, as a regressor. Finally, $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | $b_0$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | $b_6$ | $b_7$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| Model 1 | 0.0070<br>(0.36) | 0.2955*<br>(2.28) | 0.7572**<br>(4.12) | -0.309*<br>(-2.18) | -0.166<br>(-0.25) | 0.7432<br>(0.52) | 1.9471**<br>(4.03) | 1.696**<br>(-3.58) | 0.702 |

### Panel B

Estimation of the following regression:

$$MFIV_t = a_0 + a_1\, LIQUID_t + \nu_t$$
$$RVOL_t = b_0 + b_1 RVOL_{t-1} + b_2\nu_t + b_3\nu_{t-1} + b_4 LIQUID_{t-1} + b_5 LIQUID_{t-2} + b_6 ORISKS_{t-1} + b_7 ORISKS_{t-2} + u_t$$

where RVOL is the realized volatility over the remaining life of an option (corresponding to the relevant measure of implied volatility), computing using 5-minute intervals for prices, MFIV is the implied volatility computed using the Britten-Jones & Neuberger (2000) model-free implied volatility based on interpolation between traded strikes and extrapolation beyond the range of traded strikes, and $LIQ_t$ is the noise measure of Hu et al. (2012) based on mean deviations from the zero-coupon yield curve at time t. Note that this is a 2-stage procedure: in the first stage LIQUID is regressed on MFIV, and the residuals from that regression, $\nu_t$, are used as regressors in place of LIQUID so that the regressors are orthogonal by construction. We also include a lag of $\nu_t$, as a regressor. Finally, $ORISKS_t$ is the value-weighted average of the absolute value of the discount on equity closed-end funds at time t, computed as |(Price – NAV)/NAV|. All regressions are based on monthly observations between 1997 and 2008 and thus have an N of 144 observations. T-statistics are in parentheses. Notation: * & ** significant at the 5% and 1% level, respectively.

| | $b_0$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | $b_6$ | $b_7$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| Model 2 | 0.0581**<br>(2.67) | 0.2955*<br>(2.28) | 0.7908**<br>(5.76) | -0.387**<br>(-3.08) | 1.8682*<br>(2.55) | -0.608<br>(0.70) | 1.9471**<br>(4.03) | -1.696**<br>(-3.58) | 0.702 |

# Figure 1

Unique and Common Information for Forecasting Volatility
in the Presence of Past Realized Volatility