# A Corridor Fix for High-Frequency VIX: Developing Coherent Implied Volatility Measures*

Torben G. Andersen† Oleg Bondarenko‡ Maria T. Gonzalez-Perez§

February 27, 2017

**Abstract**

The VIX index is computed as a weighted average of SPX option prices over a range of strikes according to specific rules regarding market liquidity. It is explicitly designed to provide a model-free option-implied volatility measure. Using tick-by-tick observations on the underlying options, we document a substantial time variation in the coverage which the stipulated strike range affords for the distribution of future S&P 500 index prices. This produces idiosyncratic biases in the measure, distorting the time series properties of VIX. In contrast, the "Corridor Implied Volatility" index (CX) is computed from a strike range covering an "economically invariant" proportion of the future S&P 500 index values. We find the CX measure superior in filtering out noise and eliminating artificial jumps, thus providing a markedly different characterization of the high-frequency volatility dynamics.

*JEL* Classification: G13, C58

*Keywords*: VIX, Model-Free Implied Volatility, Corridor Implied Volatility, Time Series Coherence

*We are indebted to the Zell Center for Risk at the Kellogg School of Management, Northwestern University, for financial support. We thank seminar participants at the Federal Reserve Board of Governors, the SoFiE-CREATES 2010 Conference in Aarhus, the Midwest Econometrics Group 2011 Meetings, the University of Illinois at Chicago, Virginia Tech, Spot Trading, the Osaka Conference on "High-Frequency Data Analysis in Financial Markets," the Wisconsin School of Business, and the Illinois Institute of Technology for comments.

†Kellogg School of Management; Northwestern University; 2001 Sheridan Road, Evanston, IL 60208, NBER, and CREATES; e-mail: t-andersen@kellogg.northwestern.edu

‡Department of Finance (MC 168), University of Illinois at Chicago, 601 S. Morgan St., Chicago, IL 60607; e-mail: olegb@uic.edu

§Colegio Universitario de Estudios Financieros (CUNEF); Serrano Anguita, 9; 28004 Madrid, Spain; e-mail: m-gonzalezperez@cunef.edu

# 1 Introduction

The VIX volatility index disseminated by the Chicago Board Options Exchange (CBOE) has attracted much attention in recent years. It is constructed to serve as a Model-Free option-Implied return Volatility (MFIV) measure for the S&P 500 index over the coming 30 days, expressed in annualized percentage terms. The index is computed from concurrent price quotes on a wide range of out-of-the-money European put and call options written on the S&P 500 equity index, and is released in real-time by the CBOE throughout the trading day, with index values being updated about every fifteen seconds. The CBOE launched futures contracts on the VIX index in February 2004 and options on VIX futures in February 2006, enabling investors to obtain direct exposure to market volatility. The trading activity in VIX related derivatives has grown dramatically in recent years and the CBOE considers the VIX options the most successful new product launch in their history. Moreover, both the VIX futures and options have received "most innovative index product" awards at industry conferences.

The success of the VIX index and the VIX derivatives is attributable to several factors. First, there is a strong negative correlation between VIX index changes and returns on the underlying S&P 500 index. This negative relationship with market returns implies that long VIX positions can serve as a hedge against market risk and may be useful in providing diversification for portfolios that are exposed to the equity-index. On the other hand, the existence of a large volatility risk premium means that the average return on long VIX futures positions are significantly negative, implying that speculators may desire short VIX positions. Various trading strategies exploiting the latter feature have been launched under the label "volatility arbitrage" and associated performance measures are widely followed. Hence, there are natural industry constituents for the trading of VIX derivatives.

Second, the pronounced negative correlation with the overall market returns has also captured the attention of the popular press and the more specialized financial media. In fact, the VIX is often called a "fear gauge" due to the rapid spikes it displays during periods of market stress. Naturally, this moniker became even more prominent as the index soared to dramatic heights during the financial crisis of 2008-2009, reaching an intraday high of 89.5% on October 24, 2008, compared with an average value of about 19% in the period from 1990 until October 2008. In short, the VIX is now routinely referenced by the main media when characterizing the current market conditions or "investor sentiment".

Finally, the availability of a model-free market-wide measure of return volatility has proven convenient for empirical work in financial economics. It circumvents the need to specify and estimate a particular volatility model. Moreover, since it is based on option prices from a liquid market, it reflects all current relevant public information and should provide a rational signal about future return variation. One set of studies directly explores the forecast power of variation in the VIX for future realized return volatility. For example, Jiang & Tian (2005) deem the information content of the VIX volatility forecast superior to alternative implied volatility measures as well as forecasts based on historical volatility. However, a second group of studies emphasize that the VIX measure incorporates the pricing of variance risk that is embedded within option prices under the so-called risk-neutral or $Q$-measure. Hence, the wedge between regular time series forecasts for volatility, developed under the objective or $P$-measure, and the VIX, representing a risk-neutral forecast, constitutes a volatility risk premium, see, e.g., Bondarenko (2014), Bollerslev, Tauchen & Zhou (2009) and Carr & Wu (2009). The evidence favors a large, negative, and strongly time-varying variance risk premium that is negatively correlated with market returns. Interestingly, this variance risk premium has been found to carry predictability for future equity returns and to be priced in the cross-section of equity returns. Further studies document that jumps in prices and volatilities are critical for the dynamics of the variance risk premium and suggest that the premium is tied to time-varying compensation for tail risks, see Todorov (2010) and Bollerslev & Todorov (2011*b*). A third type of studies seeks to draw inference regarding the stochastic properties of market return volatility from the high-frequency behavior of the VIX. A jump in the volatility of the S&P 500 index may be hard to identify, even from high-frequency equity-index observations, while it should manifest itself directly in the VIX index. This insight motivates Todorov & Tauchen (2010) to explore the relative importance and finer probabilistic structure of jumps in high-frequency VIX series.

In a parallel development, the use of VIX style indices is moving far beyond the S&P 500 index. In recent years, the same basic methodology has been applied by leading exchanges across North America, Europe, Asia, and Australia to a broad range of asset classes and international indices. As a result, VIX measures are currently released in real-time for a host of U.S. and international equity indices as well as individual stocks, energy, metal, foreign exchange, and agricultural products. Likewise, most of these exchanges have initiated, or soon plan to initiate, trading in derivative contracts written on these volatility measures, and they are actively seeking to harmonize and standardize their procedures for measuring and trading such indices to facilitate market acceptance and depth. In short, exchange-traded products representing pure volatility bets are rapidly expanding worldwide, and the CBOE VIX methodology is emerging as the blueprint for the industry.

Given these trends, it is important to be cognizant of potential problems with the VIX measure. How robust is the measure? Does it display particular biases? Under what conditions is it likely to fail? These questions are intrinsically linked to features of the underlying market for equity-index options. This article explores the high-frequency properties of the VIX index and compares it to a number of alternative model-free volatility measures constructed from the identical SPX options. At first blush, the VIX may appear robust to market microstructure issues because it typically is computed from more than 200 distinct option mid-quotes. As such, any idiosyncratic errors will tend to cancel. Unfortunately, a range of other practical features produce systematic biases that render precise real-time measurement difficult. Many of these issues have previously been noted by Jiang & Tian (2007) who discuss how to alleviate some specific concerns. However, they do not address the lack of intertemporal consistency of the VIX index. We document that this feature induces dramatic distortions in the time series characteristics of the high-frequency VIX series and we set out to remedy them.

Initially, to establish a proper benchmark, we seek to replicate the VIX using the exact CBOE computational procedure and using the high-frequency option quotes provided by the CBOE vendor. Overall, we find a reasonable coherence between our constructed index and the official VIX. Nonetheless, there are occasional extreme deviations that have a pronounced impact on the high-frequency volatility series. The main reason is a substantial, and sometimes abrupt, time variation in the range of options exploited in the VIX computations. While we often can pinpoint the source of those shifts, they are not always consistent with our option quote record, thus driving a wedge between our replication series and the VIX. Even more importantly, however, is the fact that such sudden changes in the range of covered strikes induce artificial breaks or "jumps" in the associated volatility indices. Given these problems, we conclude that the high-frequency VIX data are not suitable for analyzing critical aspects of the volatility dynamics due to inconsistencies and noise in the time series. We also find that any alternative computation of the ideal MFIV series must confront serious obstacles in order to construct coherent index measures over time. Inevitably, the pronounced variability in the available strike coverage necessitates *ad hoc* corrections that will infuse the series with non-trivial measurement errors.

Consequently, we turn to a different methodology for volatility index construction, designed to alleviate the more critical problem with the VIX computations. Rather than using all available strikes at a given point in time, we deliberately focus on a limited strike range that, measured by an option pricing metric, covers an invariant portion of the underlying risk-neutral density. Hence, we introduce a new cutoff criterion, determined endogenously by option prices, that allows the measure to reflect the pricing of volatility across an economically equivalent fraction of the strike range, ensuring intertemporal coherence of the volatility measure. Any such measure is termed a model-free corridor implied volatility index (CX) and is closely related to the corresponding theoretical concept discussed by Carr & Madan (1998). Obviously, different corridor measures may be obtained depending on the width and positioning of the strike range. We vary the corridors moderately to explore the advantages and drawbacks of a given choice.

Our empirical results indicate that the CX indices dominate VIX in terms of providing accurate and coherent volatility measures. Specifically, the CX series contain fewer "erroneous" jumps, display stronger time series persistence, and has a higher degree of high-frequency correlation with the market returns than VIX, reflecting the improved consistency and the lower level of idiosyncratic noise of the CX measures. In addition, the measurement problems for VIX worsen during tumultuous market conditions, when the behavior of the "fear gauge" is of particular interest for traders, exchanges, regulators, and academics alike. This seriously undermines the function of the index as a real-time thermometer of market stress. Finally, we document that the quality of the end-of-day VIX value is akin to that of the intraday measure, so it is equally deficient. Hence, the distortions in the day-to-day changes in the VIX measure are not trivial.[1] However, since the (absolute) errors of the daily and the fifteen second measures are similar, but the standard deviation of daily volatility exceeds that of the high-frequency series by an order of magnitude, the relative error is much smaller when assessing the VIX at the daily level.

We conclude that our CX measures facilitate model-free identification of important underlying features of the volatility series that are useful in discriminating between alternative models and explanations for observed high-frequency asset price dynamics. Moreover, they are critical for monitoring the real-time evolution of volatility during turbulent conditions where the traditional VIX measure is especially prone to erratic behavior.

The remainder of the paper proceeds as follows. Section 2 presents the methodology underlying the model-free implied volatility measures and discusses globalization of the VIX methodology. Section 3 introduces the data used in the empirical work. Section 4 describes our replication indices RX which attempt to reproduce the high-frequency VIX. Section 5 introduce the corridor indices CX, which focus on a limited strike range that covers an invariant portion of the underlying risk-neutral density. Section 6 provides our empirical analysis of high-frequency properties of the alternative volatility indices. Finally, Section 7 concludes.

# 2 VIX Methodology

## 2.1 Model-Free Implied-Volatility and New VIX

Over-the-counter trading of volatility contracts emerged in the 1990s with liquidity building primarily in the so-called variance swaps. Despite the label, these swaps are forward contracts involving no initial or intermediary cash flows. At expiry, the long party pays a fixed premium, determined at contract initiation, and in exchange receives a payment reflecting the (random) realized variance of the underlying index. Typically, the realized variance would be measured by the cumulative sum of daily squared index returns from initiation to expiry. A critical backdrop for these developments was breakthroughs in theory demonstrating the feasibility of replicating the realized return variation of an asset over a given future horizon. This path-dependent payoff will, under general conditions, equal the payoff from a strategy involving only a static options portfolio combined with delta hedging in the underlying asset.[2] These results provide practical pricing and hedging tools necessary for sustaining market liquidity and depth. In particular, they imply that the "fair" value of a claim promising to pay the future index return variation is given by the market price of the replicating options portfolio. As such, the premium on a variance

[1] Those closing values are identical to the daily SPX option quotes available through the Wharton Research Data Services (WDRS) that form the basis for a large number of empirical option pricing studies.

[2] Initial work leading to these insights is represented by Neuberger (1994), Dupire (1993) and Dupire (1996), with the latter published in Dupire (2004). Building on these insights, Carr & Madan (1998) constructed an explicit and robust replication strategy for the pay-off on the variance swap, while Britten-Jones & Neuberger (2000) provide additional results.

swap directly values the future return variation. Moreover, since this reasoning is model-independent, the corresponding notion is referred to as the model-free implied variance, and its square-root as the *model-free implied volatility* (MFIV).

In order to facilitate the development of an exchange traded market for volatility derivatives, the CBOE, on September 22, 2003, aligned their methodology for computing the volatility index, VIX, with that of the trading community. Hence, while the "old" VIX, now denoted VXO, was based on at-the-money Black-Scholes implied volatility, the "new" VIX is designed to closely approximate the MFIV. In this manner, market participants can exploit established procedures for pricing and hedging of over-the-counter volatility contracts in dealing with exchange traded VIX products.

## 2.2 Implementation

In theory, the computation of the model-free implied variance for a given maturity requires the availability of market prices for a continuum of European-style options with strike prices spanning the support of the possible future index prices, from zero to infinity, for that specific maturity. A convenient representation of the model-free implied variance takes the form,

$$\sigma_T^2 \;=\; \frac{2\,e^{rT}}{T}\left[\int_0^F \frac{P(T,K)}{K^2}\,dK + \int_F^\infty \frac{C(T,K)}{K^2}\,dK\right] \;=\; \frac{2e^{rT}}{T}\left[\int_0^\infty \frac{Q(T,K)}{K^2}\,dK\right], \tag{1}$$

where $r$ is the (annualized) risk-free interest rate for the period $[0,T]$, as represented by the corresponding U.S. Treasury bill rate, $T$ is time-to-maturity measured in units of a year, $F$ denotes the forward price for maturity $T$, $P(T,K)$ and $C(T,K)$ are the prices for European put and call options with strike $K$ and time-to-maturity $T$, and $Q(T,K) = \min\{C(T,K),P(T,K)\}$ denotes the price of the OTM option at strike $K$. Hence, for $K < F\,(K > F)$, $Q(T,K)$ is the price of an OTM put (call). Correspondingly, we label an option with strike $K = F$, and thus $C(T,K) = P(T,K)$, an "at-the-money" (ATM) option.[3]

Of course, in practice the number of available strikes is limited so it is not feasible to compute the expression in equation (1). Instead, the CBOE approximates $\sigma_T^2$ by the following quantity,

$$\hat{\sigma}_T^2 \;=\; \underbrace{\frac{2\,e^{rT}}{T}\sum_{i=1}^{n}\frac{\Delta K_i}{K_i^2}\,Q(T,K_i)}_{\text{Discrete Approximation}} - \underbrace{\frac{1}{T}\left[\frac{F}{K_f}-1\right]^2}_{\text{Correction Term}}, \tag{2}$$

where $0 < K_1 < \cdots < K_f \le F < K_{f+1} < \cdots < K_n$ refer to the strikes included in the computation, $K_f$ denotes the first strike price available below the forward rate, $F$, where the index $f$ is a positive integer, and we assume a reasonable cross-section of options are available so that, $2 < f < n-1$. Moreover, the increments in the strike ranges are calculated as $\Delta K_1 = K_2 - K_1$, $\Delta K_n = K_n - K_{n-1}$, and for $1 < i < n$, $\Delta K_i = (K_{i+1} - K_{i-1})/2$, while the term $Q(T,K_i)$ is defined as the midpoint of the bid-ask spread for the OTM option with strike $K_i$. More specifically, $Q(T,K_i)$ is equal to the put price when $K_i < K_f$, the call price when $K_i > K_f$, and the average of the put and call prices when $K_i = K_f$. Finally, the second term in (2) reflects a correction for the discrepancy between $K_f$ and the forward price.[4]

On organized exchanges, only a few option maturity dates are quoted at any given time. The $VIX$ is defined for a fixed calendar maturity of $T_M = \frac{30}{365}$, or thirty days. The CBOE obtains the $VIX$ measure by linearly combining $\sigma_1^2$ and $\sigma_2^2$ for the two expiration dates closest to thirty calendar days with time-to-maturity of $T_1$ and $T_2$, but excluding options with less than seven calendar days to expiry. The interpolated

[3]Henceforth, we follow this standard terminology in spite of this option formally being "at-the-forward."

[4]Since $K_f$ is below $F$, the first term in (2) relies in part on higher priced in-the-money call options over the region $(K_f, F)$, resulting in an upward bias. The second negative term provides an approximate correction for this bias.

quantity is annualized and quoted in volatility units,[5]

$$VIX = 100 \times \sqrt{\left[ w_1 \left( \hat{\sigma}_1^2 T_1 \right) + w_2 \left( \hat{\sigma}_2^2 T_2 \right) \right] \times \frac{365}{30}}, \tag{3}$$

where $w_1 = \frac{T_2 - T_M}{T_2 - T_1}$ and $w_2 = \frac{T_M - T_1}{T_2 - T_1}$, so that, obviously, $w_1 + w_2 = 1$.

There are various sources of measurement error in the VIX. Jiang & Tian (2005) classify them as follows: (i) truncation errors – the minimum and maximum strikes are far from zero and infinity in equation (2); (ii) discretization errors – piecewise linear functions approximate the integrals in equation (2); (iii) "expansion" errors – a Taylor series expansion is used to approximate the log function in deriving the correction term in equation (2); and (iv) interpolation errors – linear interpolation of the maturities in equation (3).

In the remainder of this paper, our main concern is an additional, and critical, source of *idiosyncratic variation* in VIX, namely the procedure to determine the minimum and maximum strikes in computing $\hat{\sigma}^2$. The cutoff rule for the OTM options induces a *time-varying* effective strike range that produces spurious breaks in the VIX series, entirely unrelated with contemporaneous developments in the underlying return series (*artificial jumps*). We argue this problem can be greatly alleviated by implementing the Corridor Implied Volatility index (CX) as described in Andersen & Bondarenko (2007). This corridor construction explicitly controls the range $(K_1, K_n)$ in a time-consistent manner to ensure a coherent basis for computing model-free volatility.

## 2.3 Globalization of VIX

On March 16, 2011, the CBOE and Standard & Poor's announced the formation of a global "VIX Network" of exchanges with agreements regarding the use of CBOE's VIX methodology. The objective is to provide an information-sharing venue for current and potential users of the VIX methodology and to promote VIX as the global standard for measuring market volatility. At the time, agreements related to the use of VIX style measures were in place with the Australian Securities Exchange, the CME Group, Deutsche Börse, the Hong Kong Stock Exchange, the National Stock Exchange of India, Euronext LIFFE, the Taiwan Futures Exchange, and the TMX Group in Canada, among others.

Table 1 summarizes the key features behind some existing and, as of October 2014, impending volatility indices offered by these exchanges. The membership list includes most of the world's premier venues for equity and derivatives trading. It is evident that VIX related measures for equity indices are now a truly global phenomenon. Concurrently, the main U.S. exchanges have initiated VIX measures for individual stocks as well as securities linked to commodity, energy, foreign exchange and global equity exchange traded funds and futures contracts. Moreover, a number of these exchanges have introduced futures and options that trade with these volatility indices as the underlying, thus greatly expanding the set of exchange-traded "pure volatility" products.

From Table 1, it is evident that the methodology behind the index computation already is highly standardized, with exchanges largely being adhering to the CBOE VIX procedure. Some deviations arise from institutional features or liquidity concerns, while others reflect historical conventions. We briefly review the main differences here, while deferring detailed discussions until we have identified the critical aspects of the index construction.

As indicated in column two, the CBOE determines the forward rate via put-call parity for the ATM strike. This approach is accurate and robust only if the measurement errors for the quote midpoint of the ATM options are small and the quotes are current. In turn, this hinges on features like the size of the bid-ask spread and market liquidity. As an alternative, a wider set of put-call option pairs may be used to

[5] Detailed information is available at the URL: `http://www.cboe.com/micro/vix/vixwhite.pdf`.

determine the forward rate in a more robust, albeit often also less precise, manner, as noted in the Eurex regulations. Another alternative is to approximate the unobserved forward price with observed futures prices for the underlying asset, as done by the National Stock Exchange of India, the Osaka Exchange or the Moscow Exchange MICEX-RTS.

Table 1: **Volatility Indices Exploiting the VIX Methodology**

Points (1)-(4) refer to the current forward (**F**), risk-free interest rate (**r**), strikes range (**SR**) and option prices filters (**Filt**) used in computing the VIX, as reported in the CBOE's White Paper. Thus, **(1)**: the forward price **F**=$K^*_{min|C_i-P_i|} + e^{r_iT_i}[C_i - P_i]$. Mark (1.1) indicates that if more than one strike price generate the same call to put prices distance, F is equal to the average of all suitable $F_i$. Point **(2)** refers to **r** (risk-free interest rate) used to compute the VIX, the annualized yield of the U.S. Treasury bill with the closest maturity to $T_i$. According to **(3)**, the minimum and maximum strikes that define SR is decided as follows: the minimum (maximum) strike $K_{min}$ ($K_{max}$) is the OTM put (call) strike closest to $K^*$ for which the bid price of the two further strikes is equal to zero. The particular case **(3.1.)** is the same than (3), but also stoping when two consecutive zero-asks are found. Finally, $Q(K,T) = min(C(K,T),P(K,T))$ refers to ATM/OTM call and put option prices, and "min(Q)" refers to a minimum option price. **(4)** Finally, delete any remaining OTM options with zero bid prices. The acronym "n.a." indicates that information is "not available" from material provided by the exchange. "RS" denotes the "Relative Spread," computed as $RS = (A-B)/[(A+B)/2]$, where $A$ and $B$ denote ask and bid option quotes, while "MS" signifies the "maximum spread".

## CBOE / CME Group. United States

| | F | r | SR | Filt |
|---|---|---|---|---|
| **• Stock Indices:** | (1) | (2) | (3) | (4), Q |
| VIX (S&P 500), VXV (S&P 500, 3 month) , VXN (Nasdaq 100), VXD (DJIA 30), RVX (Russell 2000), VXST (9-days VIX), VSTF (Far-term VXST), VSTN (Near-Term VXST), VXMT (6-months VIX), VIN (Near term VIX), VIF (Far term VIX), VVIX (VIX volatility), VWB (Bid VIX), VWA (Ask VIX), VIXMO (S&P 500, SPXAM only), VIFMO (S&P 500, SPXAM only), VINMO (S&P 500, SPXAM only), VINXE (S&P 500, SPX Near Term Expiry), VXDEC (S&P 500 December VIX), | | | | |
| **• Interest Rates:** | (1) | (2) | (3) | (4), Q |
| TYVIX (10-years US treasury notes vol. index), SRVIX (Interest Rate Swap volatility index), SPJGBV (S&P/JPX 10-years Japanese Government Bonds) | | | | |
| **• Individual Equities:** | (1) | (2) | (3) | (4), Q |
| VXAPL (Apple), VXAZN (Amazon), VXGS (Goldman Sach), VXGOG (Google), VXIBM (IBM) | | | | |
| **• ETF (Non-U.S. Stocks):** | (1) | (2) | (3) | (4), Q |
| ⋆ <u>General</u>: VXEFA (iShares MSCI EAFE Index Fund (EFA)), VXEEM (Emerging Markets), VBEEM (VXEEM bid), VAEEM (VXEEM ask), VXFXI (China), VXEWZ (Brazil) | | | | |
| ⋆ <u>Commodities</u>: OVX (Oil), OIV (CBOE/NYMEX WTI), GVZ (Gold), VXSLV (Silver), VXXLE (Energy), VXGDX (Gold Miners), GVX (CBOE/COMEX Gold). | | | | |
| ⋆ <u>Currencies</u>: EVZ (CBOE/CME Eurocurrency ETF) | | | | |
| **• Currencies (CBOE/CME):** | (1) | (2) | (3) | (4), Q |
| EUVIX (Euro), JYVIX (Yen), BPVIX (British Pound) | | | | |
| **• Commodities (CBOE/CME):** | (1) | (2) | (3) | (4), Q |
| CIV (Corn), SIV (Soybean) | | | | |

## *Non-US* VIX Network Members

| | F | r | SR | Filt |
|---|---|---|---|---|
| **• TMX Group:** | (1) | CORRA | All | (4), Q |
| VIXC (S&P/TSX 60, Canada) | | | | |
| **• NYSE Euronext:** | (1) | (n.a.) | All | (4), Q, RS>0.5 |

Continue on next page

Table 1 – continued from previous page

| | | | | |
|---|---|---|---|---|
| VAEX (AEX, Netherlands), VFTSE (FTSE 100, UK), VCAC (CAC 40, France), VBEL (BEL 20, Belgium) | | | | |
| • **Deutsche Borse Group, Swiss Exchange:** | (1.1) | EONIA/ EURIBOR/ LIBOR | min(P) | (4), P = latest T/Q/S, A=0, MS |
| VDAX-NEW (DAX), VSTOXX (Euro Stock 50), VSMI (SMI)) | | | | |
| • **Italy (FTSE Rusell):** | (1) | OIS term | (3.1) | (4), Q, Simpson's Rl |
| FTS IVI (FTSE MIB/ FTSE 100) | | | | |
| • **Japan (OSK):** | Nikkei 225 Fut | Euroyen LIBOR (1-2M) | Avoid OTM K $\gtrless$ 6 conseqt. "invalid" OTM K | P = latest T/Q/S |
| NSAVI (Nikkei 225) | | | | |
| • **Hang Seng - Hong Kong (HKEX):** | (1) | HIBOR | $(1\pm0.2)K^{*}$ | 0<B≤ A, Q |
| HSI VIX (HSI) | | | | |
| • **Taiwan (TAIFEX):** | (1) | (2*) | (3) | (4), Q |
| TVIX (TAIEX) | | | | |
| • **Australia (ASX):** | (1) | RBA BBSW | All | (4), Q |
| S&P/ASX 200 VIX (S&P/ASX 200) | | | | |
| • **India (NSEI):** | NIFTY Fut. | NSE MIBOR | All | Q, RS>30%, (4) |
| NVIX (S&P CNX Nifty) | | | | |

Sources: http://www.cboe.com/micro/volatility/introduction.aspx, http://www.spvixviews.com/indices/, http://us.spindices.com/vix-intro/?_ga=1.89911192.2021001389.1476393943. VIX Network Members: Australian Security Exchange, CME Group, Deutsche Borse AG, Hang Seng Indexes, Japan Exchange Group, LIFFE, National Stock Exchange, Taiwan Futures Exchange, TMX Group. "The VIX Network is an association of exchanges and index providers dedicated to establishing standards that help investors understand, measure, and manage volatility. The network's members have obtained, from CBOE and S&P DJI, the rights to use the VIX methodology to calculate their own volatility indices."
($2^{*}$): the free-risk interest rate is calculated as the monthly average one-year deposit rates at the Bank of Taiwan, Taiwan Cooperative Bank, First Bank, Hua Nan Bank and Chang Hwa Bank.
Other VIX-like official volatility indices issued by Non-US / Non-VIX-member countries, are: Russia (RVI, RYSVX), South Africa (SAVI), and Korea (VKOSPI)).

A second feature is the risk-free interest rate $r$ used for discounting the option payoffs. The CBOE and CME Group interpolate these rates from U.S. Treasury bill rates, while the other exchanges rely on interbank rates, thus reflecting the costs of unsecured borrowing for major financial institutions. Since the VIX methodology is mostly applied for shorter maturities, this difference has a negligible impact, but for volatility indices covering longer maturities and during periods of financial stress with an elevated gap between interbank and treasury rates, the difference can become meaningful.

The rules governing the range of options included in the index computation are most critical. They are determined by two complementary criteria, namely the explicit conditions on the range *and* the filters invoked to eliminate faulty or excessively noisy quotes. The restrictions on the range, listed in column 4 of Table 1, display interesting variation. The CBOE applies a stopping rule centered on the ATM strike: moving into the OTM region, all options with positive bid quotes are included until two consecutive zero bid quotes are encountered, after which all further OTM options are excluded. This alleviates the noise stemming from low-priced and illiquid options, but it also induces randomness in the effective strike range. In contrast, Eurex eliminates OTM options with a mid-quote below 0.50 index points. The Hong Kong Stock Exchange uses only OTM options with exercise prices within 20% of the ATM strike, the Moscow Exchange restricts the number of OTM strikes included at each maturity to fourteen, and the Korean market establishes a sort of RND boundaries for the stock index (based on the Black Scholes parameter $d_2$ and the standardized normal CDF that results). Finally, several exchanges allow all quoted options to contribute. All these cutting-wings rules are examples of a (very inflexible) corridor implied volatility index – a notion we discuss extensively below.

The seemingly disparate rules for the option ranges are mitigated by the differences in the filters, summarized in the last column of Table 1. Exchanges that explicitly restrict the range typically only apply mild additional filtering. For example, the CBOE only adds the constraint that any remaining options with a zero bid quote are excluded, while Eurex imposes a maximum spread rule that simply enforces the existing regulations for posting valid quotes in their electronic system. On the other hand, a number of exchanges that, in principle, allow all options to enter the index computation, indirectly eliminate illiquid or low-priced options by imposing a maximum percentage spread rule. This again induces random variation in the range of options exploited for the index calculation across time. The remaining discrepancies mostly reflect institutional differences.[6]

In summary, the CBOE VIX methodology has proven extraordinarily successful. It has become the blueprint for a rapid expansion in volatility indices and products across the globe. It has captured the imagination of the public press and established itself as a gauge for market conditions. The establishment of the VIX Network solidifies the methodology's position as the worldwide standard for measuring and trading volatility across all asset classes. As such, it is imperative to study the design of the measure and understand its potential limitations for practical applications as well as academic research. We dedicate the remainder of the paper to this task.

# 3 Data

We compute model-free volatility indices from SPX option quotes obtained from the official vendor of CBOE data, Market Data Express, or MDX. The data stem from an electronic feed supplied by two of the lead market makers in the SPX options. The series are based on MDR (Market Data Retrieval) quotes captured by CBOE's internal system. The underlying tick-by-tick data cover the period June 2, 2008 – June 30, 2010, encompassing 520 individual trading days, and contain a huge number of quotes.[7] For the two maturities involved in the construction of the implied volatility measures, we have an average of about 213,000 (147,000) OTM put (call) option quotes per day, or around $52\,(33\frac{1}{2})$ million OTM put (call) option quotes in total across the full sample.

We construct fifteen second series for each individual option using the "previous tick" method, implying that we retain the last available quotes prior to the end of each 15 second interval throughout the active trading day, from 8:30:15 to 15:15:00 CT.[8] If no new quote arrives in a 15 second interval, the last available quote prior to the interval is retained. Hence, in the case of inactivity, the quotes may become stale. Consequently, for a variety of analysis, we impose a limit on the duration of widespread inactivity as described below. In addition, as a benchmark, we will on occasion exploit the corresponding intraday quotes for the S&P 500 futures from the CME Group.

Following Andersen, Bondarenko & Gonzalez-Perez (2015), a few filters are applied to guard against data errors. First, we avoid excessive staleness by deeming an option missing if the quote time-stamp precedes the computation of the volatility index by more than five minutes. If quotes for nearby strikes are available, the VIX computation interpolates the option value for this strike during the calculation via

[6] For example, some exchanges stipulate that the ask quote cannot be lower than the bid quote. Other exchanges enforce this constraint within their electronic quoting system and do not add this requirement to their explicit filtering procedure.

[7] The size of the full data set is around 263 GB, and it is constructed from 525 daily CSV files of an average size of about 0.5 GB each. The files include tick-by-tick quotes and trades for options on the S&P 500 index across all active maturities and strikes. From our empirical analysis, we discard five Holidays with only partial trading: July 3, 2008; November 28, 2008; December 24, 2008; November 27, 2009; and December 24, 2009.

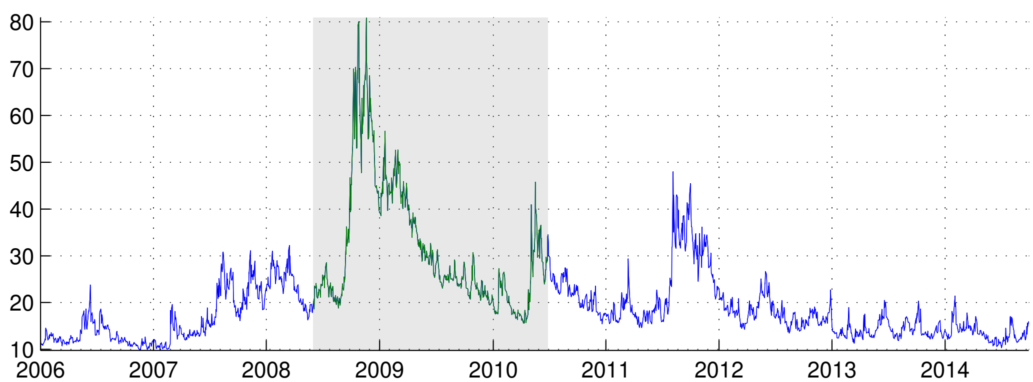
[8] For each 15 second cross-section, we exploit an average of 66 (39) strikes for OTM put (call) options for the first maturity, and an average of 69 (43) strikes for OTM put (call) options at the second maturity.

the formula in equation (2), so this does not create a gap in the volatility series. Second, if an entire block of adjacent and relevant OTM option prices have not been updated for five minutes, we deem the index itself "not available" (n.a.) and eliminate this time period to avoid errors due to "systemic staleness." This condition is essential for avoiding artificial jumps in the volatility index when an entire set of stale option quotes are updated simultaneously after a temporary malfunction of the quote dissemination system. Third, we impose a maximum threshold for the degree of no-arbitrage violations implied by the option mid-quotes, and we declare the volatility index n.a. whenever this "non-convexity" threshold is exceeded. We lose close to 1.2% of the 15 second observations due to our filters, with about half of the missing values stemming from the non-systemic staleness requirement and half from the non-convexity condition.[9]

The corresponding fifteen second series for VIX is obtained from different sources although all quotes originate with the CBOE. The first data set, covering June 2008 to September 2008, provides the VIX values at fifteen second intervals and was obtained directly from the exchange. The second source is MDR-VIX which contains tick-by-tick quotes of VIX options for October 2008 to April 2009. This source indicates the value of the underlying VIX index whenever a VIX option quote is recorded, so a complete fifteen second series may be extracted. Finally, the May 2009 – June 2010 period is again covered by the fifteen second VIX series, but now secured via MDX.

We apply only a very mild filter to the intraday VIX series: we remove observations that fall outside the daily high-low range for VIX. Since this range is reported at the end of trading, it may reflect the correction of errors in the real-time VIX values disseminated during the day. We denote this very lightly filtered series VIX$^*$. Figure 1 shows the evolution of the daily VIX closing value over recent years, with the shaded region denoting the sample period exploited in this article.

Figure 1: **Daily Closing Value for VIX, January 2006 – August 2014**. Our sample period, June 2, 2008 – June 30, 2010, is indicated in green and is shaded.




## 4 Practical Issues with the Computation of VIX

This section describes in detail the steps necessary to replicate the high-frequency VIX, characterizes sources of bias in the CBOE methodology, studies why it is not feasible to replicate the index at a zero matching error, and characterizes a new significant source of error of VIX – the random time-variation of the effective strike ranges used to compute the index.

[9]The procedures for implementing these filters are detailed in the Appendix.

### 4.1 First attempt to replicate VIX: RX0 and RX1 indices

We first seek to replicate the official VIX index from the underlying SPX options using the exact CBOE methodology. The initial step involves the computation of $\hat{\sigma}^2$ for each relevant maturity. However, the forward price, $F$, and the strike price just below the forward price, $K_f$, in equation (2) are not directly observable. The CBOE determines these basic variables from the available option quotes in three steps. First, for each maturity, they identify the strike price $K^*$ for which the distance between the quoted midpoints of the call and put prices is minimal. This strike is then used to compute the "implied" forward according to put-call parity, $F^* = K^* + e^{rT} \times [C(K^*,T) - P(K^*,T)]$. Finally, $K_f$ is determined as the first available strike price below $F^*$. We refer to the VIX index constructed from (2) and (3), using $F^*$ and the associated $K_f$, as our first "Replication indeX", or RX0. This index constitutes our direct equivalent of the CBOE VIX, computed according to official rules and using official data.

Our analysis has revealed that the way the CBOE finds the ATM strike price $K_f$ entails a certain lack of robustness, arising from the use of only a single ($K^*$) put-call option price pair to infer $F^*$ and $K_f$. Occasional problems, like a temporary gap in the quote updating for a subset of options or a basic recording error, can cause the minimum distance between the call and put prices to, erroneously, appear at a point far from the true ATM strike. In turn, this can generate a large bias in the VIX, and may even render its computation infeasible.

Motivated by this finding, we propose a more robust procedure for determining the implied forward. In a first step, for a given maturity, we identify the set of strikes for which the discrepancy between the price of the put and call is below \$25. Secondly, for each of these put-call option pairs, we compute the implied forward price. Third, we designate the "robust implied forward" to be the median of this set of implied forward prices. Using a robust statistic defined over multiple option pairs prunes the implied forward series of major erroneous outliers. On the other hand, the wider set may include less liquid options with large relative pricing errors. Consequently, as our final step, we retain the original $F^*$ value, rather than the robust forward, unless the two values differ substantially. Following diagnostic checks, we chose a relative threshold of 0.5%.[10] Hence, if the two implied forwards are close, we stick with $F^*$, but if they differ by more than 0.5%, we instead adopt the robust implied forward. This procedure generates an implied forward price, $F$, that typically coincides with $F^*$, but deviates whenever $F^*$ may contain a sizeable error. To illustrate the difference between the two forwards, Figure 2 plots the ratio $F^*/F$. Two vertical lines delimitate the 0.5% absolute difference between the two forwards, above which we replace $F^*$ with $F$ in the computation of the volatility index. The ratio falls below the 0.995 threshold in 0.055% of times and is above the 1.005 threshold in only 0.002% of times. Many of these cases occur during the first five minutes of trading.

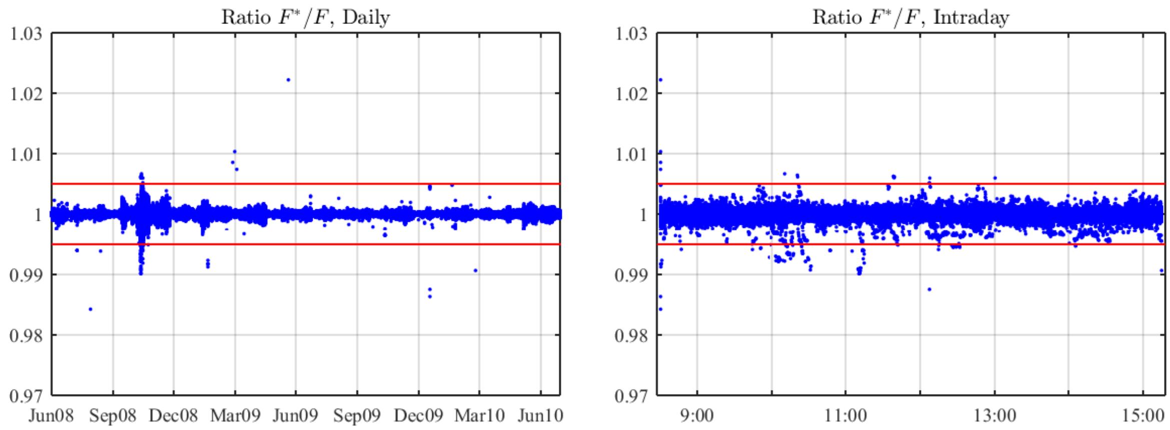
We denote the VIX replication index, computed from $F$, as RX1. Importantly, it retains the feature that, apart from the risk-free interest rate, it is computed exclusively from CBOE option prices, as is the case for the official VIX. It is interesting to note that this approach is consistent with the spirit of the forward price computation employed by NYSE Euronext, as indicated in Table 1.[11] Throughout our analysis we include both RX0 and RX1 as candidate VIX measures although the conclusions generally are very similar. We also rely on the robust forward price $F$ when computing all other indices to be explained later.

Figure 3 depicts the evolution of the volatility indices RX1 and VIX* on October 6, 2008. The left panel shows that RX1 is able to replicate VIX* very well throughout the day. Unfortunately, the impression is somewhat deceptive. The right panel of Figure 3 zooms in on a shorter period within the day.

[10] For a level of the S&P 500 around 1000, the threshold is about 5 which is the same order of magnitude as the gap between the strike prices.

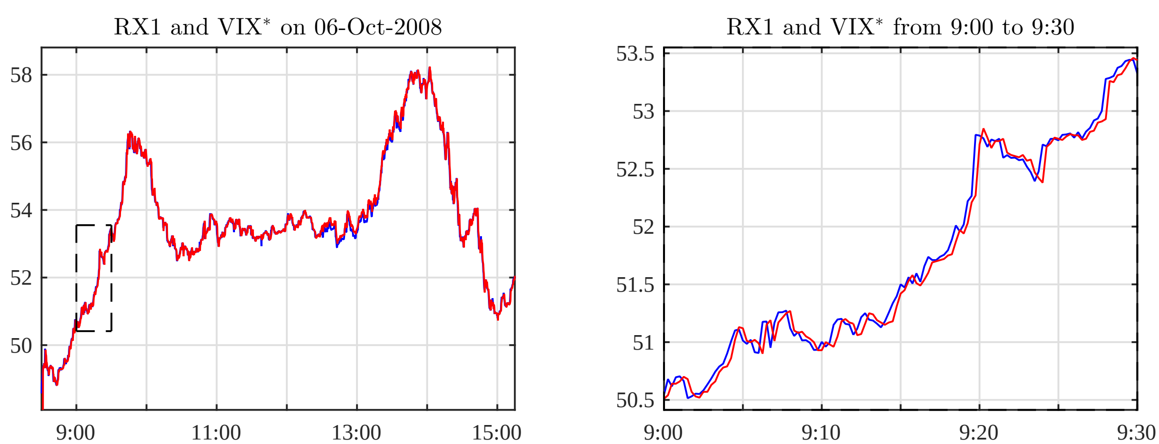
[11] Some markets impose exceptions to the general CBOE rule to compute the forward, but these refer to the existence of more than one strike with the same minimum difference between call and put options. For example, Eurex states in its white paper for VDAX: "If a clear ($F^*$) minimum does not exist, the average value of the relevant forward prices will be used instead".

Figure 2: **Ratio of CBOE Forward to Robust Forward**. The top panel shows ratio $F^*/F$ intraday (for each of 1620 15-second intervals, there are 520 observations). The bottom panel shows the same ratio over the whole sample (for each of 520 trading days, there are 1620 observations). Red horizontal lines indicate 0.5% difference between forward values.




Close inspection reveals that the real-time VIX lags our RX1 series by about 15-30 seconds. The delay is not unique to this period or trading day, but is observed consistently throughout the sample. It is likely caused by capacity constraints of the CBOE server which processes data for a large set of derivatives contracts simultaneously.[12] Since the reporting delay varies over time in ways that, given the nature of the CBOE system, are not observable and cannot be reconstructed *ex post*, it is infeasible to construct a record of the actual quotes used for the real-time VIX computation. Consequently, it is impossible to match the high-frequency VIX values perfectly, even using the official methodology. Moreover, the discrepancies arising from this delay can be non-trivial, even within a fairly calm trading environment like this. For example, just after 9:00, the RX1 measure rises quickly from 50.5% to about 53.5%. Given the delay, this will mechanically create a gap relative to VIX*, even if the two indices are constructed from the identical option prices. Moreover, this type of deviation, due purely to a dissemination lag, becomes more prominent in times of rapidly shifting volatility, or high volatility of volatility, when the VIX index is of particular interest as a gauge of evolving market conditions.

Figure 3: **RX1 and VIX* on October 6, 2008**. RX1 (blue) and VIX* (red) are shown for October 6, 2008. The right panel zooms in on the 30-minute period after 9:00, indicated by the black dashed rectange on the left panel.




[12]This explanation arose from conversations with John Hiatt, Director of Research at CBOE.

While the above issues raise some concerns regarding the integrity of the real-time VIX series, they are probably not – in and of themselves – sufficient to cause major reservations. First, large discrepancies due to the implied forward price are rare and readily handled. Second, some real-time delays in the dissemination are to be expected, and if market movements are reflected correctly with only a small delay the associated distortions of the statistical properties of the index will be minor. Third, such problems constitute much less of a problem for assessing the volatility dynamics at lower frequencies. In particular, the typical shift in volatility over 15-30 seconds is minuscule compared to those observed over daily horizons, so the relative errors due to the delay are trivial for daily data. This does not imply that such discrepancies are immaterial, but rather that their importance hinges on the intended use of the volatility measure.

Our main concerns about the VIX index reside elsewhere and run deeper as we explain below.

### 4.2 All-available-strikes Replication Index RX2

In principle, all (available) options should be exploited in computing a model-free volatility index and hence also the VIX. However, this strategy faces some practical problems. First, far OTM options are near worthless and trade infrequently, if at all. Second, partially as a result, the relative bid-ask spread is very high for far OTM options and even the midpoint quotes may be poor indicators of underlying value. As discussed previously, the CBOE adapts to this feature by invoking a specific cutoff rule: moving away from the forward price, once two consecutive strikes with zero bid quotes are encountered, all further OTM options are discarded. Hence, market liquidity and pricing jointly determines the cutoff level. The rule has intuitive appeal. For example, as volatility rises and option (bid) quotes increase, then additional, and newly valuable, options are included in the VIX calculation. As a result, the computation should generally include all options with non-trivial (bid) valuations and it should avoid the noise and distortions associated with far OTM options which, in theory, have only a limited impact on the overall measure. On the other hand, this introduces the potential for random variation in the strike range that may complicate the comparison of VIX measures across time and distort the time series properties of the index.

We rely on the concept of an *effective strike range*, or ER, to gauge the coverage of options in the VIX computation. Formally, the effective range is defined as,

$$ER \,=\, [LT, RT] \,=\, \left[ \frac{\ln(K_1/F)}{\hat{\sigma}_{BS}\sqrt{T}}, \frac{\ln(K_n/F)}{\hat{\sigma}_{BS}\sqrt{T}} \right], \tag{4}$$

where $LT$ ($RT$) is the left (right) truncation point corresponding to the lowest (highest) strike $K_1$ ($K_n$) and $\hat{\sigma}_{BS}$ is an ATM implied Black-Scholes volatility measure, obtained from a linear interpolation of four Black-Scholes implied volatilities, extracted from options with strike prices just above and below the forward price for the two maturities closest to 30 calendar days. Using the definition in (4), we calculate the effective ranges $ER_1$ and $ER_2$ corresponding to the two option maturities $T_1$ and $T_2$ used in the VIX computation. We then linearly interpolate $ER_1$ and $ER_2$ to obtain the weighted effective range $ER$, which corresponds to the constant maturity of 30 calendar days. To make our illustrations less cluttered, we focus on the weighted effective range $ER$, as opposed to its components $ER_1$ and $ER_2$.

The effective range reflects the degree of coverage afforded by the given set of options, controlling for the concurrent level of volatility. It allows us to assess whether the CBOE stopping rule produces a coherent strike range across time for the computation of VIX. If we instead ignore the stopping rule, and exploit all options with positive bid quotes, we obtain a broader range, reflecting the *maximum* coverage offered by the prevailing market liquidity. We denote this broader index RX2. It should provide an upper bound on the VIX values reported by the CBOE. Moreover, whenever there are no positive bid quotes for OTM options beyond the CBOE truncation point, RX1 and RX2 coincide. We also note that, as indicated in Table 1, the use of all (valid) option quotes is consistent with the conventions adopted by NYSE Euronext and the main exchanges of Australia and India.

Figure 4: **RX1, RX2 and VIX$^*$ on November 7, 2008**. The top left panel depicts RX1 and RX2, while the top right panel adds VIX$^*$ (red) to the other two volatility indices. The bottom left panel displays the effective strike ranges used by RX1 and RX2 throughout the day. The bottom right panel plots the ratio of RX1 and RX2 to VIX$^*$. All measures for RX1 are shown in blue and for RX2 in green.

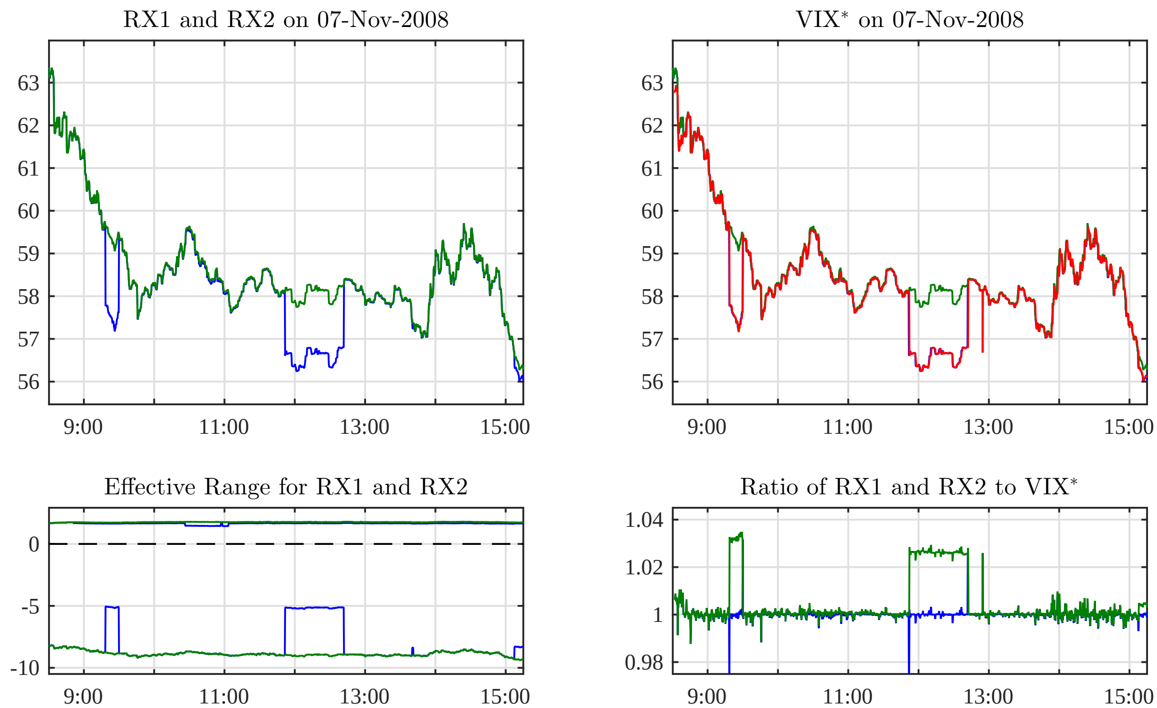
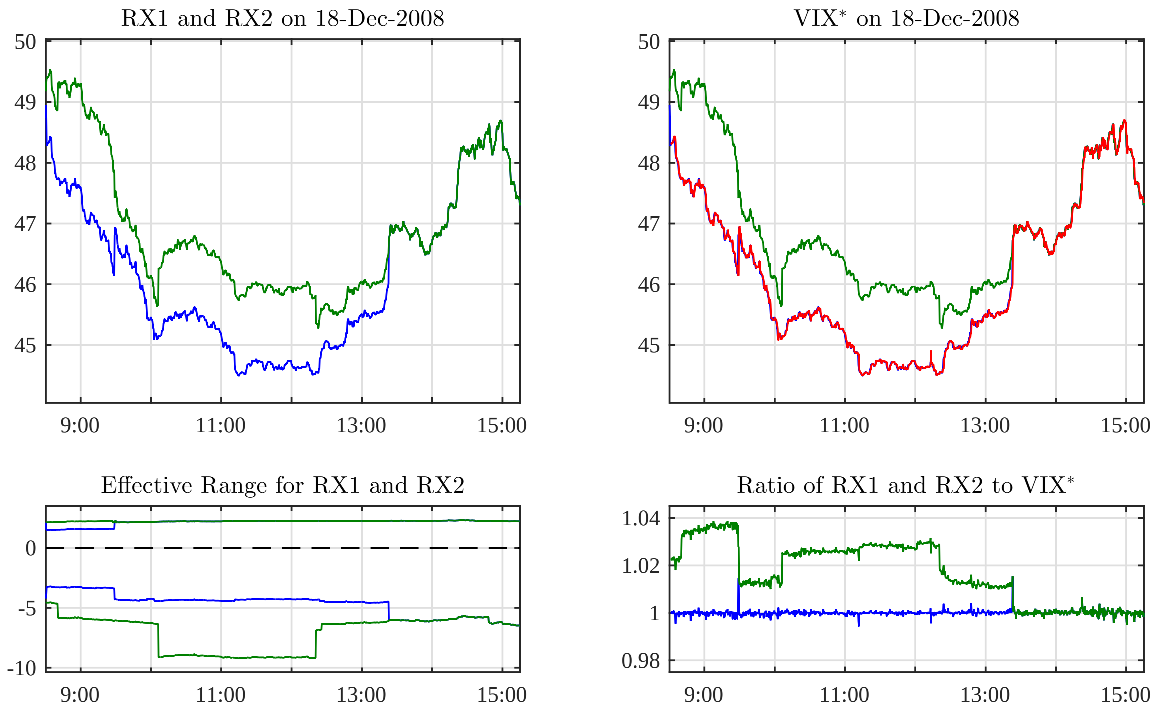



Figure 4 illustrates why the strike range underlying the VIX computation may be of concern. On November 7, 2008, RX1 replicates VIX almost perfectly throughout the day. RX2 and RX1 are also mostly the same, except for two separate periods. The first period starts at around 9:18, when RX1 and VIX abruptly decrease from 59.7 to 57.8 in 15 seconds, which corresponds to more than 3% of the index value. This move is then completely reversed at around 9:30, when RX1 and VIX jump up and converge again with RX2. The explanation for the two offsetting moves is apparent from the left bottom panel of Figure 4, where we notice dramatic concurrent shifts in the lower part of the strike range (deep OTM put options) used for computing RX1. In contrast, the RX2 measure is computed from a near invariant strike range and, as a result, its evolution, almost exclusively, reflects the changes in the underlying volatility. The same scenario is repeated during the second special period, around 11:52-12:42. As earlier, RX1 and VIX jump down from RX2 and then later jump back, at exactly the same points when the effective coverage of RX1 shrinks and then expands. Overall, during this trading day, the four big moves in RX1 and VIX represent what we call "dubious" or "artificial" jumps. Because these moves are very large and happen over a very short periods, they will be classified as jumps by essentially every existing jump-detection method. And yet, the apparent "jumps" are completely driven by the abrupt shifts in the effective strike range of RX1 and not by any coincident change in volatlity.

It is clear from Figure 4 that the cutoff rule for the strike range determined by the CBOE does not always produce the intended result. Using the full range of available strikes would have been preferable for the particular day explored above. When inspecting other days, RX2 often appears like an improvement to RX1 and VIX. Unfortunately, RX2 is too plagued by related problems and it does not completely

Figure 5: **RX1, RX2 and VIX$^*$ on December 18, 2008**. The top left panel depicts RX1 and RX2, while the top right panel adds VIX$^*$ (red) to the other two volatility indices. The bottom left panel displays the weighted effective strike ranges used by RX1 and RX2 throughout the day. The bottom right panel plots the ratio of RX1 and RX2 to VIX$^*$. All measures for RX1 are shown in blue and for RX2 in green.



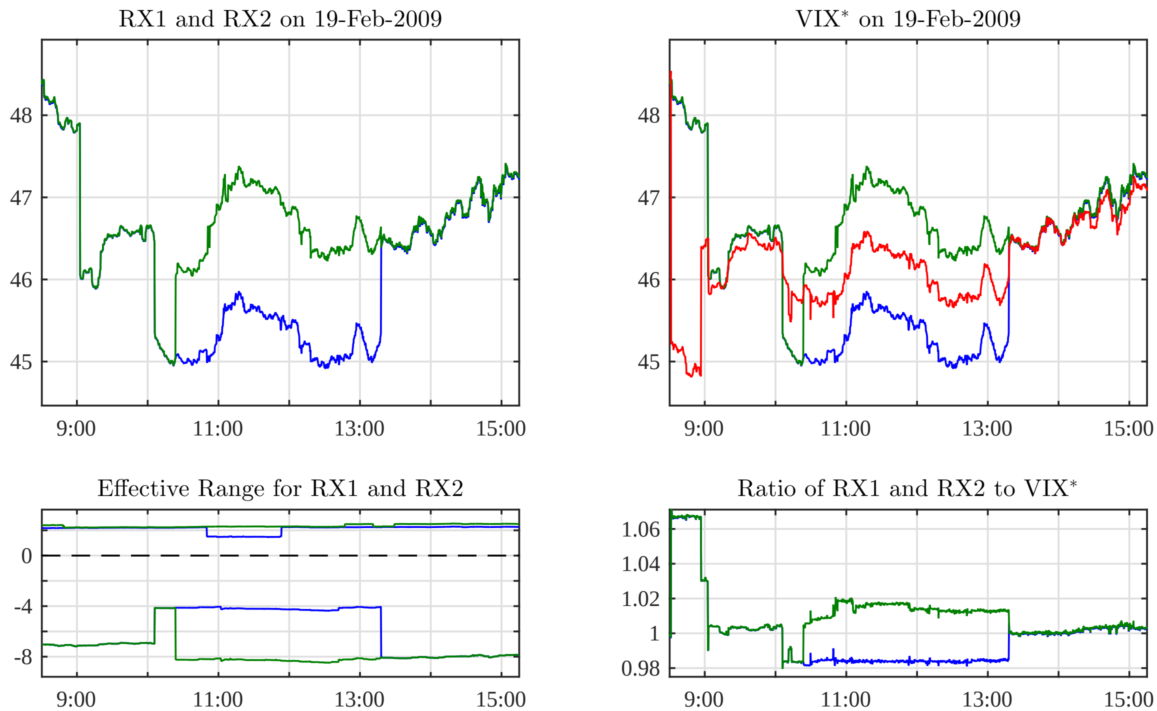

solve the effective range issue.

Consider, for example, a trading day illustrated by Figure 5. On December 18, 2008, VIX is successfully replicatedl by RX1 throughout the day. RX1 and VIX have two artificial jumps up, at around 9:29 and 13:23, when the effective strike range for RX1 abruptly expands. After the second jump up, RX1 and VIX coincide with RX2. However, RX2 also exhibits two artificial jumps, which largerly offset each other. The first (up) jump occurs at around 10:06 and the second (down) jump occurs at around 12:21. The jumps are caused by the sharp expansion in the lower part of the strike range (deep OTM put options) used to compute RX2. This range expansion is responsible for more than 1% inflation in the index's value.

Finally, Figure 6 illustrates a trading day with a "kitchen sink" of issues. On February 19, 2008, sometimes VIX is reasonably well approximated by both RX1 and RX2 and sometimes by neither one. In the latter case, the approximation errors are as high as 7%. There exist periods when RX1 and RX2 are very close to each other and periods when they are far apart. There are big artificial jumps in RX1 (at around 10:06 and 13:18) and RX2 (at around 10:06 and 10:24). There is also a large *true* jump in both RX1 and RX2 at around 9:03. At this point, the indices jump down from 47.9 to 46.0 (more than 4%), while their effective ranges remain largely unchanged. Interestingly enought, around the time when RX1 and RX2 have a true negative jump, the official VIX actually jumps up.

It should be pointed out that the problems illustrated in Figures 4-6 are very common and are observed across many other days. Those are not some special, or atypical cases.

Figure 6: **RX1, RX2 and VIX* on February 19, 2008**. The top left panel depicts RX1 and RX2, while the top right panel adds VIX* (red) to the other two volatility indices. The bottom left panel displays the weighted effective strike ranges used by RX1 and RX2 throughout the day. The bottom right panel plots the ratio of RX1 and RX2 to VIX*. All measures for RX1 are shown in blue and for RX2 in green.




## 4.3 Coherence between VIX and RX Indices

We have argued that perfect replication of the high-frequency VIX series is infeasible due to a time-varying dissemination delay at the CBOE. However, as long as the VIX and RX2 indices are based on the same methodology and a similar set of underlying option quotes, the average discrepancies should be minor, even if they can be non-trivial during high volatility periods. Figure 7 depicts over time the Mean Absolute Percentage Error (MAPE) for the high-frequency RX1 and RX2 series relative to VIX*, where

$$MAPE = 100 \times \frac{1}{n} \sum_{t=1}^{n} \left| \frac{VIX_t^* - RX_t}{VIX_t^*} \right|, \qquad RX = RX1,\ RX2.$$

The left panel of the figure displays MAPE averaged for each month. As expected, the deviations are larger in the depth of the financial crisis. Nonetheless, the overall coherence is good, with an average discrepancy of about 0.35% for RX2 and 0.19% for RX1. Moreover, it is less than 0.1% for RX1 over the last eleven months of the sample, when market conditions were less turbulent.

The right panel of Figure 7 depicts the intraday pattern for MAPE. In this case, the absolute percentage errors are averaged across days for each 15-second interval. The deviations are quite large at the market open and they generally decline over the course of the trading hours. Outside the first 5 minutes, there are two additional spikes: at 9:00, the time when many macroeconomic announcements are made, and at 15:00, when the equity market closes. Both cases are associated with sharp increases in volatility and overall trading activity.

In Table 2, we compute the fraction of time the indices RX1 and RX2 coincide with VIX*, using three relative deviations (0.25%, 0.50% and 0.75%) as the criterion for satisfactory replication. The first

Figure 7: **Mean Absolute Percentage Approximation Error**. MAPE for the RX1 (blue) and RX2 (green) relative to VIX*. The left panel shows monthly MAPE, with the straight lines indicating full sample averages. The right panel shows intraday MAPE.

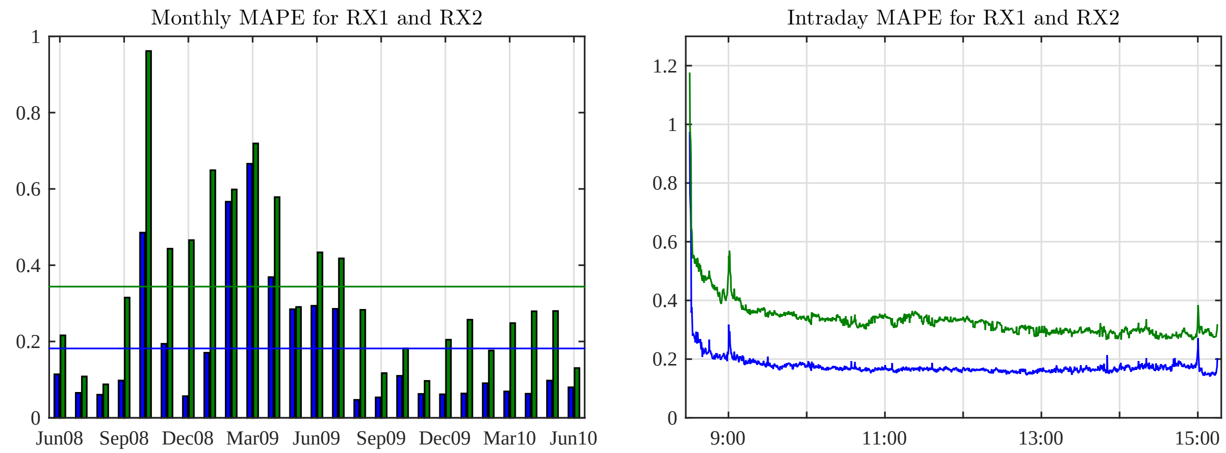



Table 2: **Coherence between the VIX, RX1, and RX2.** The table reports the fraction (%) of cases for which the VIX* is matched by RX1, RX2, or both at a given threshold for relative accuracy (0.25%, 0.5% and 0.75%). Statistics are reported separately for intraday (15 sec) and end-of-day (EOD) cases. We only consider intraday data from 8:35 CT. Percentage are computed over data with non-missing values for all three indices (the percentages of missing data re 1.18%, and 1.54% for 15-seconds and EOD series, respectively).

| | Accuracy = 0.25% | | | Accuracy = 0.50% | | | Accuracy = 0.75% | | |
|---|---|---|---|---|---|---|---|---|---|
| | RX1 | RX2 | Both | RX1 | RX2 | Both | RX1 | RX2 | Both |
| 15 sec | 80.5 | 64.6 | 63.6 | 91.0 | 81.2 | 80.7 | 95.5 | 88.6 | 88.2 |
| EOD | 79.7 | 63.5 | 61.3 | 91.2 | 79.5 | 77.9 | 95.5 | 90.0 | 89.1 |

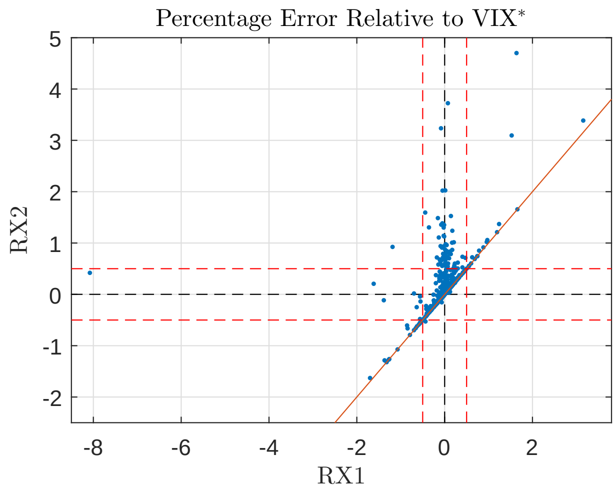
row in Table 2 reflects the fraction of the 15-second VIX observations that are replicated within the given threshold for accuracy, while the second row provides the corresponding figure for the end-of-day value of the VIX. The latter measure constitutes the daily (closing) value of the VIX and is available from the CBOE website.

Consider the intraday case first. Consistent with the fact that RX1 is designed to explicitly mimic the CBOE cutoff rule, this index achieves the highest degree of replication of 80.5% (91.0%, 95.5%) at the 0.25% (0.50%, 0.75%) threshold. More interestingly, at the 0.5% threshold, 80.7% of the time VIX* values are compatible with both RX indices. In the remaining 19.3% of the cases, we can provide a close replication of the VIX* index by RX2 in 0.5% of the time and by RX1 in 10.3% of the time. This still leaves 8.5% of the observed VIX* values unaccounted for.[13] It is also striking that the percentages are very similar for the end-of-day VIX* values. Hence, our findings for the intraday frequency carry over

[13] We performed a few robustness checks to further gauge these relationships. Most importantly, there is a time-varying delay, averaging between 15 and 30 seconds, in the reporting of the official VIX* after SPX quotes are recorded. Consequently, any abrupt change in volatility will induce an artificial, short-lived deviation between the VIX* and RX series (the latter inherits the SPX time stamps). Hence, we explored the match between the VIX* and 15- or 30-second lagged RX series. While the fractions of matched observations increased, the effect was marginal. Thus, our count of unexplained deviations may be slightly inflated relative to the "true" discrepancies, but this is a minor bias.

to the daily VIX measure! Figure 8 shows the distribution of end-of-day approximation errors for RX1 and RX2. Again, most of the time RX1 provides a close replication to VIX*. There is a small number of cases when RX2, but not RX1, replicates VIX*. There are also quite a few cases when RX1 and RX2 are very close to each other, but they are far from VIX*.

Figure 8: **Approximation Errors for RX1 and RX2**. This figure shows scatter plot of end-of-day percentage errors for RX1 and RX2. Percentage errors are relative to VIX*. The red dashed lines indicate 0.5% bounds, while the solid green line is the 45-degree line.




In summary, considering the inevitable discrepancies arising from the CBOE and the data dissemination delay, we are for the most part able to replicate the high frequency VIX index with reasonable precision. Nonetheless, we do identify troublesome features. Most importantly, the VIX, on occasion, displays large discrete changes solely due to abrupt shifts in the range of underlying option strikes (see Figures 4 and 5). The RX2 measure is less susceptible to such errant movements and may be a more reliable volatility gauge than the VIX series. However, since RX2 uses all available strikes it will be impacted by the relatively large bid-ask spread, or noise, associated with far OTM options. Moreover, the effective strike range for RX2 may also, occasionally, vary significantly with overall option market liquidity, thus generating the same type of problem as observed for VIX* and RX1. We can only assess the extent of this problem by explicitly controlling for the effective range used in the computation of the index.

# 5 Corridor Implied Volatility Indices CX

To assess how variation in the strike range impacts the volatility measures, we turn to the concept of corridor implied volatility, introduced by Carr & Madan (1998) and explored empirically in Andersen & Bondarenko (2007) and Andersen, Bondarenko & Gonzalez-Perez (2015). The point is, *a priori*, to fix the range of strikes at a level that provides broad coverage but avoids excessive extrapolation of noisy or non-existing quotes for far OTM options. Since it operates with a fairly narrow range, the Corridor Index, or CX, will usually be lower than RX2. However, the invariant strike coverage ensures that the measure is coherent in the time series dimension and alleviates the variation induced by idiosyncratic shifts in the effective range.

The main issue is how to define an economically invariant portion of the strike range which ensures that the associated implied volatility measures are compatible over time. There are a number of ways to measure moneyness or the degree of "tailness" that have been exploited in the literature. One common approach is to measure moneyness relative to the forward price via the distance measured in Black-Scholes (BS) ATM implied volatilities. Another popular approach is to measure moneyness in terms of option Black-Scholes delta, which conveniently ranges from 0 to 1. Both of these approaches, however, rely on a model-dependent, and clearly misspecified, assumption of constant volatility and log-normality. In particular, the Black-Scholes volatility fails to adapt suitably to variation in the shape of the risk-neutral distribution (RND) caused by shifts in the skewness or by asymmetric variation in the tail thickness. Another common approach is to rely on the percentiles of the RND.[14] In principle, this approach is model-free, but, unfortunately, it also has significant drawbacks. First, RND percentile are not centered on the forward price. This is particularly troublesome for equities due to the pronounced skew in the RND. In fact, across our sample, the forward price is located anywhere between the 38th and 47th percentiles of the distribution with a mean around the 43th. Second, RND percentiles are not reliable indicators of the importance of the associated strikes, as time-varying tail-thickness has a first order impact on option prices. Finally, estimation of the RND is computationally non-trivial, requires auxiliary assumptions, and depends on chosen method. In practical terms, it would be challenging for an exchange to advocate an index whose computation is non-transparent and hard to replicate.[15]

In response to these shortcomings, we adopt an alternative approach developed in Andersen & Bondarenko (2007) and also implemented in Andersen, Bondarenko & Gonzalez-Perez (2015). In this approach, the truncation barriers are based on the following ratio statistic:

$$R(K) \;=\; \frac{P(K)}{P(K)+C(K)}, \tag{5}$$

where $P(K)$ and $C(K)$ denote put and call option prices for strike $K$ and the same time to maturity $T$. Despite its simplicity, this ratio statistic possesses a number of desirable features. First, the statistic depends on option prices only and is trivial to compute over the strike range for which active quotes are available. Second, $R(K)$ is akin to a cumulative density function in that it is strictly increasing in K over $(0,\infty)$, with $R(0)=0$ and $R(\infty)=1$. Third, the ATM put and call prices are theoretically identical, so that $R(F)=0.5$, where the forward price $F$ equals the mean of RND, that is, $F=E^Q[F_T]$. This implies that the forward price, $F$, separating the strike range into OTM put and call regions, also represents the median of the $R(K)$ function. Consequently, the percentiles of the ratio function conveniently center the strike range on the focal point for the computation of model-free implied volatility. This is particularly useful when considering a set of nested CX measures as we do below.

Let $K_q$ define the $q$ "percentile" of the $R$ function:

$$K_q \;=\; R^{-1}(q) \qquad \text{for any} \qquad q\in[0,1]. \tag{6}$$

It is natural to define the truncation points for the implied volatility computation via symmetric percentiles $[K_q, K_{(1-q)}]$, so that the truncation points reflect the relative importance of the right and left tails for option pricing in a consistent manner across time. In this paper, we consider three corridor measures, CX1– CX3 which correspond to the choice of $q=0.01, 0.03, 0.05$. These choices of $q$ produce wide, moderate, and narrow corridors, with CX2 representing our preferred benchmark. The corridor indices CX are then computed exactly as the RX indices, except that the barriers are given by $B_L=K_q$ and $B_H=K_{(1-q)}$.

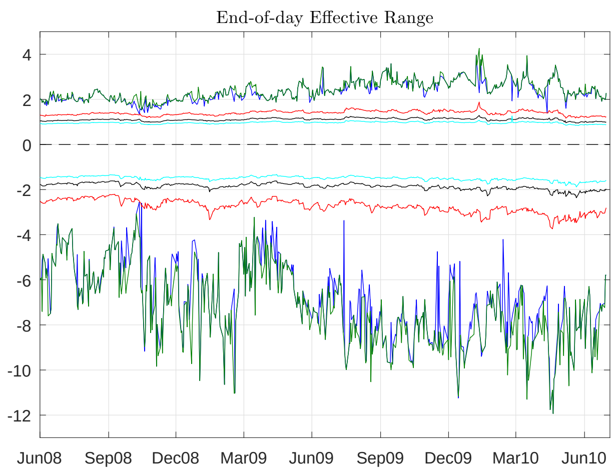
[14] The convention of defining corridors via the percentiles of the RND was adopted by Andersen & Bondarenko (2007).

[15] As noted in Table 1, the Hong Kong Stock Exchange uses a corridor defined in terms of the moneyness of the strikes. This corridor is unresponsive to shifts in volatility and does not adapt to variation in the shape of the RND at all. As a result, it will cover a highly varying proportion of the underlying RND over time, and produce an intertemporally incoherent implied volatility measure.

### 5.1 Assessing the Stability of the Corridor Strike Range

We have established that shifts in the effective strike range can be an important source of idiosyncratic jumps in implied volatility measures constructed along the lines of the VIX. This motivates our introduction of the CX indices, which compute volatility measures from a consistent span of option strikes over time. This section compares the stability of the effective strike ranges associated with alternative MFIV indices. Figure 9 depicts the end-of-day coverage for various indices measured in terms of implied ATM Black-Scholes volatilities.

Figure 9: **Effective Range for Volatility Indices.** End-of-day weighted effective strike range for RX2 (green), RX1 (blue), CX1 (red), CX2 (black), and CX3 (cyan).



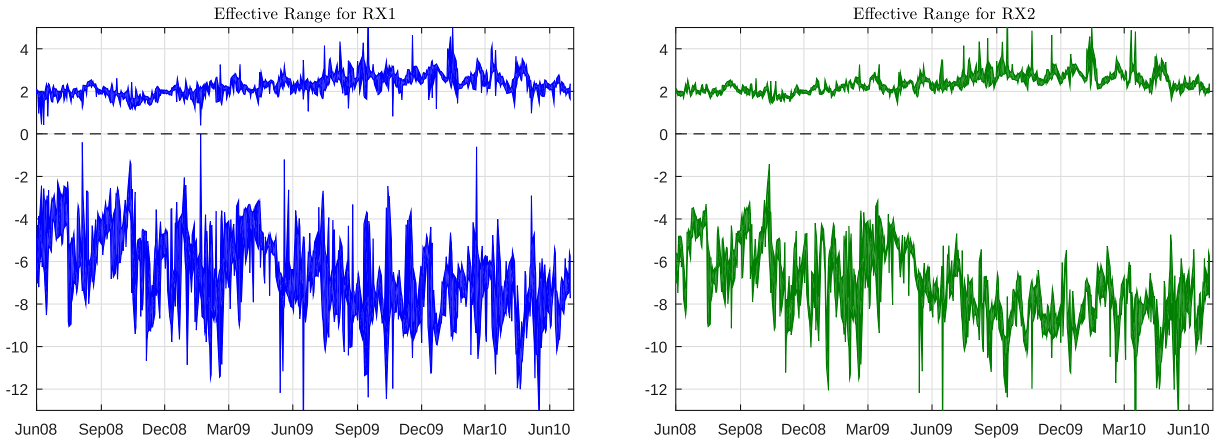

First, we observe the pronounced asymmetry in the coverage of OTM puts and calls, with the put range typically being more than double the call range. This underscores the relative importance of the OTM put versus call options for pricing volatility. Second, we note that the variation in the strike coverage of the RX measures is dramatic, especially in contrast to the stability attained by the CX measures. This suggests that corridor indices will provide superior intertemporal coherence in the volatility measures. Third, comparing the individual RX measures, we note that RX1 and RX2 have largely matching outliers, but RX1 sports more frequent inlier observations which arise from an intermediate truncation of the RX1 strike coverage. Since such "premature" truncation has a significant impact on implied volatility measures because closer-to-the-money options carry a higher price tag, the RX2 series may be more reliable than RX1 in terms of capturing fluctuations in market volatility. Nonetheless, the strike coverage for both RX1 and RX2 are extraordinarily variable, with the downside truncation point $ER^L$ ranging across $[-11.9, -3.1]$ and $[-11.9, -2.6]$, respectively. In comparison, the variation in the downside truncation level for CX1-CX3 is much smaller, falling within the intervals $[-3.7, -2.2]$, $[-2.4, -1.6]$, and $[-1.9, -1.3]$, respectively.[16]

Figure 9 focuses on only one point during each day, the market close. Therefore, it does not reveal how much strike ranges vary intraday. Figure 10 depicts this information for RX1 and RX2. Specifically, for each index, this figure shows the whole interval of values that strike range takes on a given trading day. Again, the strike range is relatively more stable for RX2 than RX1, but both exhibit enormous variation within many days. For example, focusing on the left truncation point $LT$ (deep OTM put options), there are 93 (24) instances when this cuttoff changed by more than 4 sigmas within a single day for RX1 (RX2). Similarly, there are 26 (6) separate trading days when $LT$ changed by more than 6 sigmas for

[16] Of course, if the degree of truncation is measured not in Black-Scholes sigmas, but instead in terms of percentiles of the model-free $R$ function, then indices CX1-CX3 have, by construction, constant coverages.

RX1 (RX2). Clearly, variations in the strike range of such magnitude are likely to induce considerable distortions in RX indices, unrelated to shifts in option prices.

Figure 10: **Effective Range for RX1 and RX2**. For each trading day, filled areas show minimum and maximum for the left and right truncations point $LT$ and $RT$.




To explore whether the variation in the effective range has a systematic and economically meaningful effect on volatility indices, we follow two different approaches.

## 5.2 OLS Regression on Effective Strike Range

The first approach is based on a simple regression

$$\frac{X}{CX2} - 1 = \beta_0 + \beta_1 \cdot LT(X) + \beta_2 \cdot RT(X) + \varepsilon, \qquad X = RX1,\ RX2,\ VIX^*, \tag{7}$$

where the dependent variable captures the value of volatility index $X$ relative to the benchmark index CX2 and the explanatory variables $LT(X)$ and $RT(X)$ are the left and right truncation points of the effective strike range. We estimate this regression separately for end-of-day (EOD) series and 15-second series and report in Table 3. Consider the EOD case first in Panel A. The results for RX1 and RX2 indices indicate that the coefficients for both explanatory variables are highly statistically significant and have correct signs. The effect of the $LT$ (OTM puts) is much stronger than that of $RT$ (OTM calls). Overall, the adjusted $R^2$ shows that the effective strike range has very strong explanatory power for the observed variation of RX1 and RX2 relative to CX2.

Althought the results for our replication indices RX1 and RX2 are very compelling, ideally, we would like to estimate a similar regression but for the official VIX, which is employed by virtually all academic studies. However, in this case, we do not observe the exact strike range that CBOE uses to compute VIX. Therefore, we approximate unobservable truncation points $LT(VIX^*)$ and $RT(VIX^*)$ with observable $LT(RX1)$ and $RT(RX1)$, since RX1 typically replicates VIX* quite well – see Table 2. We estimate the regression two times: (1) for all observations, and (2) for only those observations where RX1 matches VIX* to within 0.5% accuracy. The results are reported in the last two rows of Panel A in Table 3. The regression $R^2$ remains very high for both specifications, despite the fact that the effective range of VIX* is now measured with some noise. As expected, the fit improves for the second specification, where VIX* is matched more closely by RX1.

The last row of Panel A indicates that when $LT(RX1)$ decreases by one sigma (i.e., when the effective range widens), VIX* will be on average higher by 0.93%. As discussed earlier, there are many days when

Table 3: **OLS Regression on Effective Strike Range.** The table reports the results of OLS regression in (7) for end-of-day (EOD) and 15-second series. For VIX* regressions in the last two rows, true $LT(VIX^*)$ and $RT(VIX^*)$ are unobservable and are approximated by $LT(RX1)$ and $RT(RX1)$. Specification (1) uses all observations, while specification (2) uses only those observations where RX1 matches VIX* to within 0.5% accuracy. The robust $t$-statistics reflect Newey-West HAC standard errors based on 5 lags (EOD) or 40 lags (15-sec).

**Panel A: End-of-Day**

| | $\beta_0 \cdot 10^2$ | $\beta_1 \cdot 10^2$ | $\beta_2 \cdot 10^2$ | $t(\beta_0)$ | $t(\beta_1)$ | $t(\beta_2)$ | $\bar{R}^2$ |
|---|---|---|---|---|---|---|---|
| RX1 | 1.00 | -0.90 | 0.77 | 1.78 | -15.4 | 3.6 | 68.7 |
| RX2 | 0.80 | -0.89 | 0.81 | 1.24 | -15.3 | 3.1 | 71.0 |
| VIX* (1) | 1.28 | -0.87 | 0.76 | 2.20 | -13.7 | 3.7 | 66.2 |
| VIX* (2) | 0.92 | -0.93 | 0.70 | 1.59 | -15.8 | 3.3 | 70.8 |

**Panel B: 15-second**

| | $\beta_0 \cdot 10^2$ | $\beta_1 \cdot 10^2$ | $\beta_2 \cdot 10^2$ | $t(\beta_0)$ | $t(\beta_1)$ | $t(\beta_2)$ | $\bar{R}^2$ |
|---|---|---|---|---|---|---|---|
| RX1 | 1.42 | -0.89 | 0.61 | 30.09 | -168.1 | 28.8 | 68.1 |
| RX2 | 1.07 | -0.88 | 0.72 | 20.17 | -172.4 | 30.8 | 67.7 |
| VIX* (1) | 1.56 | -0.85 | 0.70 | 31.72 | -143.6 | 32.2 | 64.1 |
| VIX* (2) | 1.25 | -0.90 | 0.64 | 25.78 | -156.3 | 29.1 | 67.5 |

the range of RX1 experiences shifts in $LT$ of multiple sigmas. These purely idiosyncratic shifts in the strike range, unrelated to shifts in the underlying option prices, induce economically large distortions in VIX*.

The results for the intraday case are reported in Panel B of Table 3 and they are generally consistent with the EOD case. The regression $R^2$ statistics are now somewhat lower and the slopes $\beta_1$ and $\beta_1$ are slightly closer to zero. Since the number of 15-second observations is enormous (almost 1 million), $t$-statistics are very high.

### 5.3 Simulated Indices SX1 and SX2

Regression in (7) is simple and parsimonious, but it has limitations. First, it assumes a linear relationship between the volatility index and its effective range, which is clearly counterfactual. When $LT(X)$ changes from -2 to -3 or from -10 to -11, in both cases the shift is one sigma, but the effect on the volatility index is very different. Second, the explanatory variables in (7) correspond to the *weighted* strike range, obtained by interpolating the effective ranges $ER_1$ and $ER_2$ for the two option maturities $T_1$ and $T_2$. In situations where ranges $ER_1$ and $ER_2$ differ substantially, linear interpolation can distort the relationship between the weighted volatility index and its weighted strike range.

In this subsection, we compliment the regression approach by introducing the so-called *simulated* indices. At any time, index RX1 (RX2) reflects both the fundamental volatility as reflected in option prices and the stirke range used in the index calculation. We would like to disentagle the two effects and proceed as follows. We select a representative day $d$ in our sample and compute option implied volatilities for available strikes. These volatilities are interpolated and extrapolated for a wide range of strikes, to create a continuous volatility curve. We then assume that this implied volatility curve remains constant for our whole period. That is, for any day, we replace true option prices with simulated ones, corresponding to the fixed implied volatility curve. Except for prices, other option characteristics (maturities $T_1$ and $T_2$, effective strike ranges $ER_1$ and $ER_2$, etc.) are kept unchanged. This way, we obtain *simulated* indices SX1 and SX2, which are computed exactly as RX1 and RX2, except for using simulated prices. By construction, variation in simulated indices SX1 and SX2 is due to variation in effective strike range only. Indices SX1 and SX2 show what RX1 and RX2 would be if implied volatility skew (the fundamental volatility) remained invariant and equal to that on the reference day. Notice also that for the corridor indices CX1-CX3, their simulated counterparts would be always constant, since implied volatilities, option prices, the *R*-function, and the corridors would be fixed.

Figure 11: **End-of-day Simulated Indices SX1 and SX2**. The dashed line shows the sample mean.

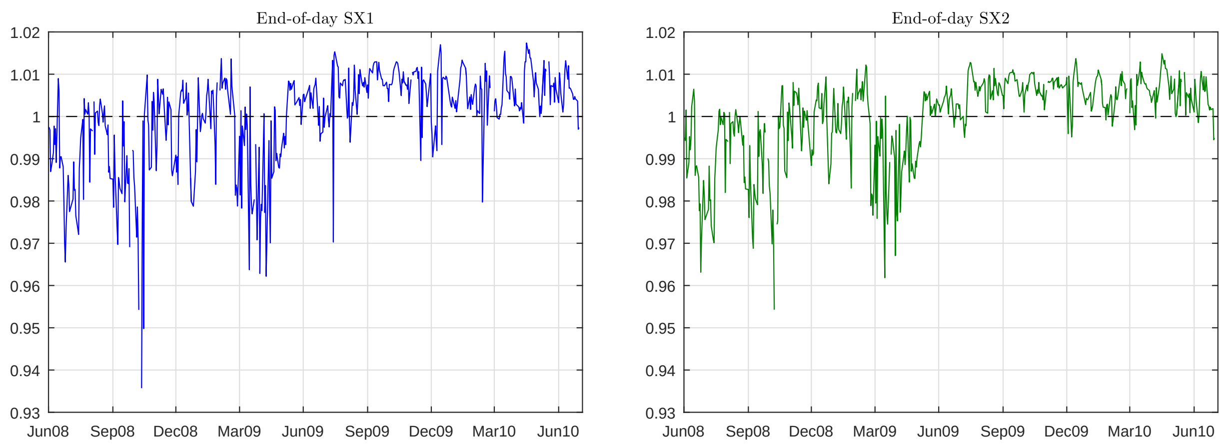
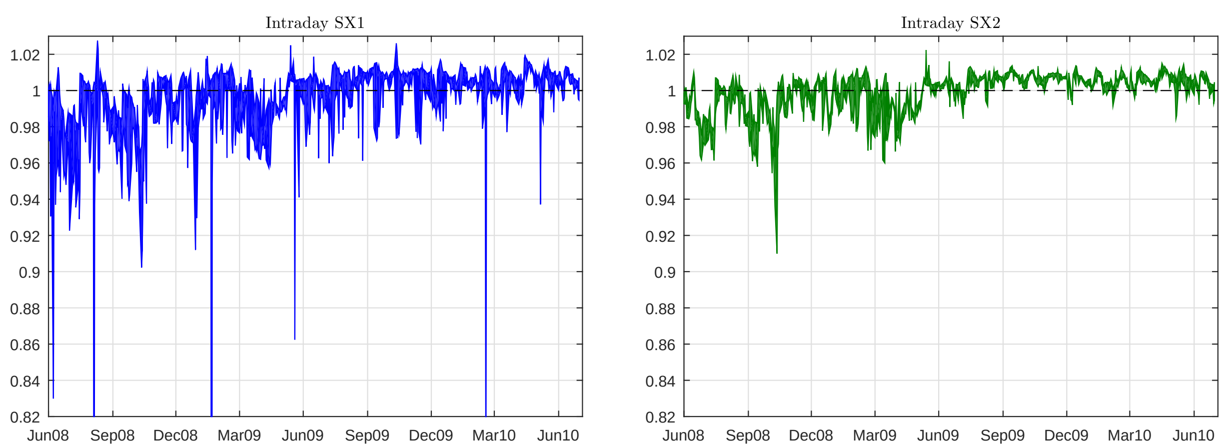



Figure 12: **Intraday Range for SX1 and SX2**. For each trading day, filled areas show the intraday range for the simulated indices SX1 and SX2. The dashed line shows the sample mean.




To construct a representative implied volatility curve, we take all days for which $T_1$ is exactly 30 days (thus, avoiding the need to model two maturities separately), compute their EOD ATM implied volatilities and select the median.[17] The selected day is August 20, 2009, when the ATM implied volatility is 21.9. Since the level of simulated indices is somewhat arbitrary, to simplify interpretations, we normalize each index to be one on average. Figure 11 depicts the end-of-day levels of simulated indices SX1 and SX2. Both indices exhibit considerable variation from day to day. For example, the difference between the highest and lowest values of SX1 is more than 7%. Figure 12 shows the whole interval of values that SX1 and SX2 take on a given trading day. Consistent with our previous findings for the effective strike

[17] As a robustness check, we also consider days for which ATM implied volatility is in 10th and 90th percentiles.

range, we observe that SX1 is much less stable than SX1. In particular, there are 28 (5) instances when SX1 changed by more than 5% (10%) within a single day.

# 6 Properties of High-Frequency Volatility Indices

Table 4 provides the summary statistics for the alternative volatility measures. The most noteworthy aspect of Table 4 is the huge variation in the kurtosis statistic for the volatility changes, or returns, in Panel B. While the sample kurtosis for the CX series are sizeable, falling in the range of 21 to 23, the values for RX1, RX2 and VIX* are rather imposing at 55, 90 and 63, while it attains a truly outsized value of 4874 for VIX.

Table 4: **Summary Statistics, 1-min Frequency**. $\rho_i$ indicates the $i^{\text{th}}$ autocorrelation coefficient, while $\rho_{i,j}$ denotes the average of the $i^{\text{th}}$ through the $j^{\text{th}}$ autocorrelation coefficients.

**Panel A: Volatility Index Levels**

| | $RX1$ | $RX2$ | $CX1$ | $CX2$ | $CX3$ | $VIX^*$ | $VIX$ |
|---|---|---|---|---|---|---|---|
| Mean | 31.58 | 31.64 | 30.27 | 29.03 | 28.10 | 31.61 | 31.62 |
| Std Dev | 13.28 | 13.33 | 12.78 | 12.31 | 11.94 | 13.28 | 13.29 |
| Skewness | 1.33 | 1.32 | 1.32 | 1.32 | 1.32 | 1.32 | 1.32 |
| Kurtosis | 4.21 | 4.19 | 4.20 | 4.20 | 4.19 | 4.18 | 4.17 |
| % Missing | 1.19 | 1.18 | 1.17 | 1.17 | 1.17 | 1.04 | 0.98 |

**Panel B: Volatility Index Returns**

| | $RX1$ | $RX2$ | $CX1$ | $CX2$ | $CX3$ | $VIX^*$ | $VIX$ |
|---|---|---|---|---|---|---|---|
| Mean $\times 10^4$ | -0.11 | -0.16 | -0.12 | -0.14 | -0.13 | -0.10 | -0.10 |
| Std Dev $\times 10^4$ | 21.66 | 20.39 | 21.09 | 21.72 | 21.95 | 21.45 | 29.23 |
| Skewness | 0.11 | -0.57 | 0.34 | 0.33 | 0.35 | 0.06 | -0.44 |
| Kurtosis | 54.79 | 90.15 | 23.04 | 21.74 | 22.59 | 63.07 | 4873.72 |
| $\rho_1 \times 10^2$ | 5.24 | 6.22 | 5.35 | 4.71 | 4.10 | 5.27 | 3.59 |
| $\rho_{2,5} \times 10^2$ | 2.63 | 2.95 | 2.67 | 2.27 | 2.17 | 2.79 | 2.48 |
| $\rho_{6,21} \times 10^2$ | 1.00 | 1.07 | 0.98 | 0.88 | 0.85 | 0.94 | 0.90 |
| $\rho_{21,45} \times 10^2$ | 0.03 | 0.04 | 0.03 | 0.00 | -0.00 | 0.01 | 0.01 |
| % Missing | 1.46 | 1.45 | 1.43 | 1.43 | 1.43 | 1.31 | 1.23 |

Notice also that mild filtering reduces the kurtosis of VIX* dramatically relative to VIX, suggesting that erroneous data are an important source of the inflated statistics. Overall, these findings are consistent with the view that all series, bar perhaps the CX indices, suffer from sizeable artificial outliers. Another point of note is the relatively small skewness statistics of the volatility returns. There is little evidence of any pronounced asymmetry in volatility changes, contradicting the idea that (high-frequency) volatility jumps only, or predominantly, are positive. Finally, we observe that there is positive serial correlation in the volatility returns, but the effect is largely dissipated after five minutes and entirely gone after twenty minutes. If the VIX measures were error-free and direct trading on the VIX index was feasible, one would expect the VIX returns to display minimal short-term autocorrelation as competing agents eliminate simple predictable return patterns. Hence, the short term return persistence is likely indicative

Table 5: **Correlations, 1-min Frequency**

**Panel A: Volatility Index Levels**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* |
|---|---|---|---|---|---|---|
| $RX1$ | 1.00000 | | | | | |
| $RX2$ | 0.99980 | 1.00000 | | | | |
| $CX1$ | 0.99953 | 0.99956 | 1.00000 | | | |
| $CX2$ | 0.99914 | 0.99918 | 0.99988 | 1.00000 | | |
| $CX3$ | 0.99884 | 0.99889 | 0.99972 | 0.99996 | 1.00000 | |
| $VIX^*$ | 0.99993 | 0.99974 | 0.99952 | 0.99912 | 0.99882 | 1.00000 |

**Panel B: Volatility Index Returns**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* |
|---|---|---|---|---|---|---|
| $RX1$ | 1.000 | | | | | |
| $RX2$ | 0.871 | 1.000 | | | | |
| $CX1$ | 0.863 | 0.910 | 1.000 | | | |
| $CX2$ | 0.856 | 0.895 | 0.964 | 1.000 | | |
| $CX3$ | 0.850 | 0.887 | 0.958 | 0.969 | 1.000 | |
| $VIX^*$ | 0.672 | 0.607 | 0.622 | 0.614 | 0.611 | 1.000 |

of some staleness in the underlying option quotes.[18]

In spite of pronounced discrepancies across the return series, the indices are near identical in their portrayal of the general volatility level, as is apparent from Panel A of Table 4. Except from the fact that the CX indices, as a consequence of deliberate truncation, represent slightly down-scaled versions of model-free implied volatility, the statistics are strikingly similar across the board, e.g., the kurtosis measures all take values close to 4.2. Hence, apart from a slight reduction in the mean level and standard deviation, the CX measures capture the identical features of overall volatility as the VIX and RX indices. This translates into an extraordinarily high degree of correlation between the index levels, as documented in Panel A of Table 5. In contrast, correlations between volatility index returns are considerably lower, as documented in Panel A of Table 5. In particular, VIX* has correlations between 0.61 and 0.62 with CX indices. It achieves the highest correlation of 0.67 with RX1.

## 6.1 Volatility Jumps

As in Andersen, Bondarenko & Gonzalez-Perez (2015), we proceed nonparametrically by defining large moves relative to a robust measure of concurrent volatility. We first obtain a robust estimate of the volatility for each index series across every trading day, accounting both for shifts in volatility across days and the pronounced intraday volatility pattern.[19] Exploiting this robust standard deviation measure, we sort the one-minute returns for each volatility measure into a set of mutually exclusive size categories. Table 6 tabulates the findings.

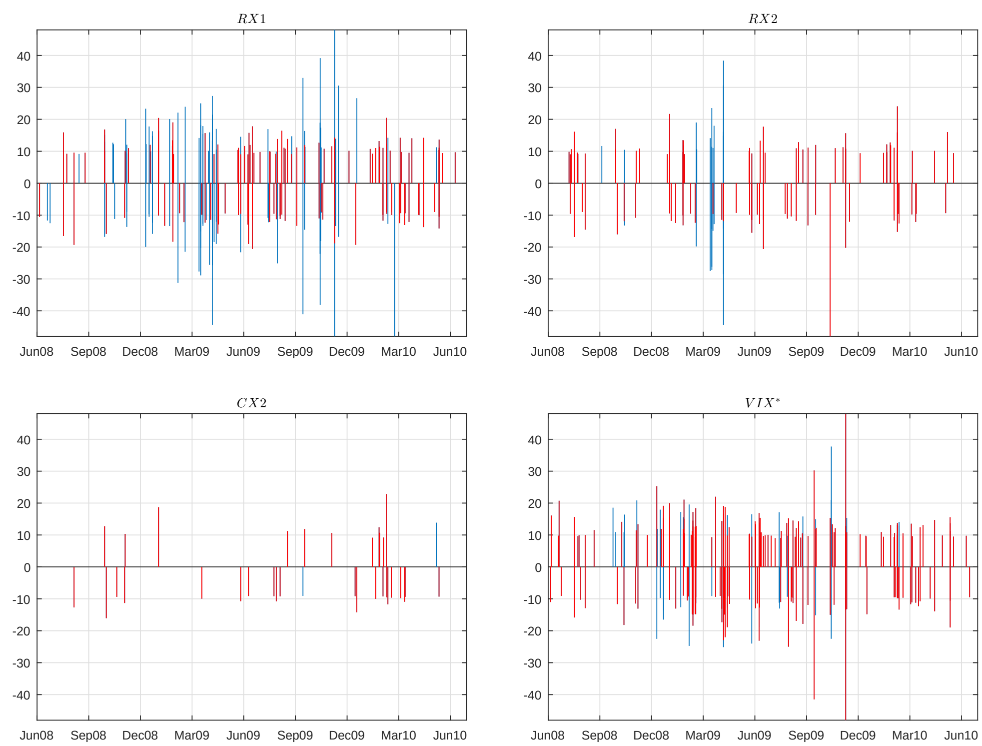
[18]Trading on the VIX is feasible, e.g., via the VIX futures contracts offered by the CBOE. However, the VIX "cash index" explored here is not directly traded.

[19]A description of this procedure is given in Appendix B. Moreover, we have confirmed that the qualitative results are robust to a number of alternative ways of estimating intraday volatility.

Table 6: **Distribution of Extreme Returns ("Jumps"), 1-min Frequency**. Range is measured in multiples of the (robust) standard deviation.

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* | VIX |
|---|---|---|---|---|---|---|---|
| $(-\infty, -30)$ | 6 | 2 | 0 | 0 | 0 | 2 | 12 |
| $(-30, -15)$ | 27 | 11 | 3 | 1 | 1 | 21 | 36 |
| $(-15, -9)$ | 66 | 43 | 16 | 21 | 21 | 70 | 99 |
| $(-9, -6)$ | 239 | 194 | 141 | 127 | 118 | 260 | 263 |
| $(-6, -4)$ | 718 | 682 | 636 | 614 | 607 | 727 | 727 |
| $(4, 6)$ | 726 | 655 | 602 | 590 | 604 | 717 | 719 |
| $(6, 9)$ | 248 | 208 | 162 | 139 | 126 | 245 | 248 |
| $(9, 15)$ | 79 | 48 | 16 | 10 | 9 | 92 | 124 |
| $(15, 30)$ | 33 | 12 | 2 | 2 | 3 | 34 | 48 |
| $(30, \infty)$ | 4 | 2 | 0 | 0 | 0 | 3 | 10 |

Figure 13: **Jumps over Time.** This figure depicts time-series of returns which fall into the LJ group (*Large jumps*). Positive jumps are blue and negative jumps are green. Returns are measured in multiples of sigma.

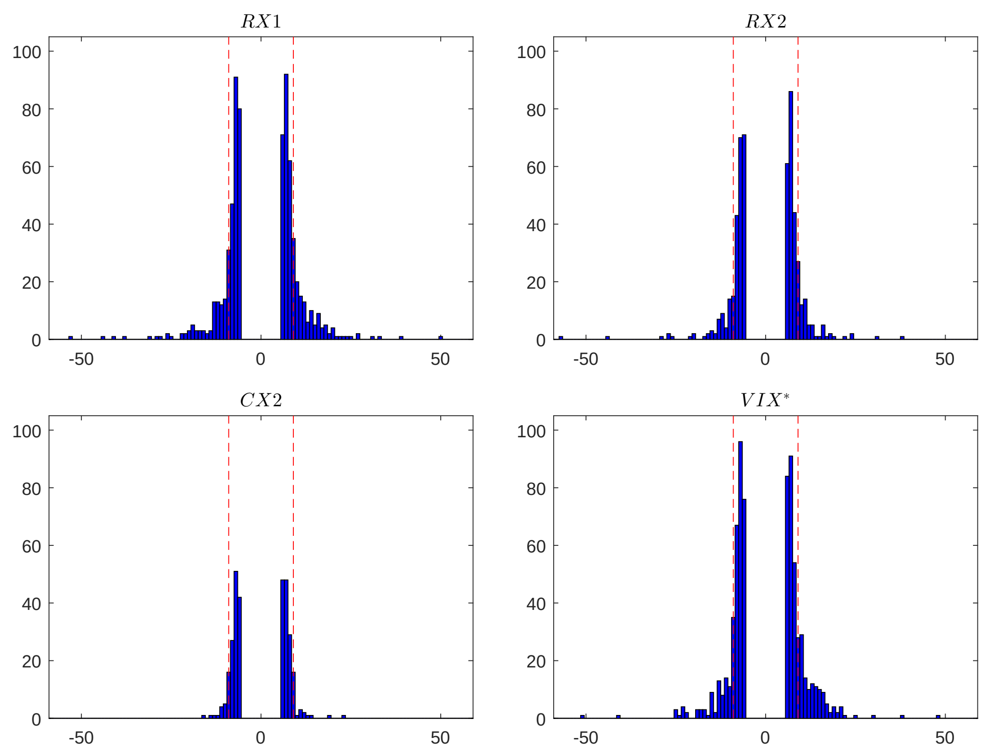



The table reveals a startling discrepancy in the number of large moves across the alternative indices, irrespective of the threshold adopted for identifying "jumps." For example, including moves beyond nine standard deviations on either the upside or downside, we find 307 returns in VIX, 217 in VIX*, 205 in RX1, and 114 in RX2. In contrast, the numbers for CX1-CX3 are 37, 34, and 34 returns, respectively. Given that the sample covers 25 months, VIX* and VIX display, on average, about 2 and 3 large moves per week, while CX2 displays, on average, only one large move every 3 weeks. Even more telling is the discrepancy in the number of extremely large moves of more than 15 standard deviations. The CX2 measure has 3 such moves, while VIX* and VIX have 60 and 106. Put differently, about 95% (97%)

of all extreme moves in VIX* (VIX) are likely to be spurious. Similarly, the number of moves of more than 15 standard deviations for RX1 and RX2 are many times more than for CX2. This is obviously anomalous: any genuinely large shift in volatility should manifest itself in a significant elevation of option prices across a broad range of strikes and should be reflected in all broadly based model-free volatility indices. This suggests that various sources of measurement error, including the documented idiosyncratic variation in the strike range, are sufficiently commonplace that they severely inflate the outlier count for VIX and RX indices.[20]

A second intriguing finding is the near symmetry of positive and negative "jumps" within each size category. Part of the explanation may be that misclassified jumps naturally revert, as an unusual widening of the strike range, say, is followed by a sudden return to the usual range. However, even for measures less prone to this type of error, we observe a near symmetric distribution of positive and negative jumps. This suggests that volatility frequently moves significantly in either direction. The evidence of numerous *negative* volatility jumps is inconsistent with many popular parametric specifications of asset price dynamics allowing only for positive jumps in volatility.

Figure 14: **Jump Histograms.** This figure depicts distributions of one-minute returns which fall into one of the two groups: SJ (*Small Jumps*) and LJ (*Large jumps*). Red dashed lines separate SJ and LJ. Returns are measured in multiples of sigma.




For expositional purposes, we classify 1-minute returns into three groups, *regular moves* (RM), *small jumps* (SM), and *large jumps* (LJ), which are defined as returns with absolute value less than 6 standard deviations, or sigmas, between 6 and 9 sigmas, and more than 9 sigmas, respectively. Figure 13 displays LJ returns for various volatility indices across the full sample. Positive jumps are blue, and negative ones are green. Compared to CX2, large jumps in indices RX1, RX2, and are much more frequent and they are typically larger in magnitudes. Most of those jumps are artificial. Figure 14 displays return

[20]The indices in Table 6 contain a slightly different number of one-minute observations due to the difference in the filters applied to the series. The qualitative results are unaffected by restricting the sample to the smaller set of one-minute returns that are common to the series. These findings are available upon request.

histograms for the same four volatility indices. It show only those returns which fall into either SJ or LJ groups. Again, the contrast between CX2 and the other three indices is very striking, especially in the extreme tails.

Figure 15: **Negative Jump of CX2 on October 1, 2008.** The top panels depict levels of CX2 and S&P 500, while the bottom panels depict the corresponding 1-minute returns, measured in multiples of sigma. At 12:44, CX2 has a big negative jump of -16.1 sigmas, while S&P 500 has a positive jump of 15.9 sigmas.

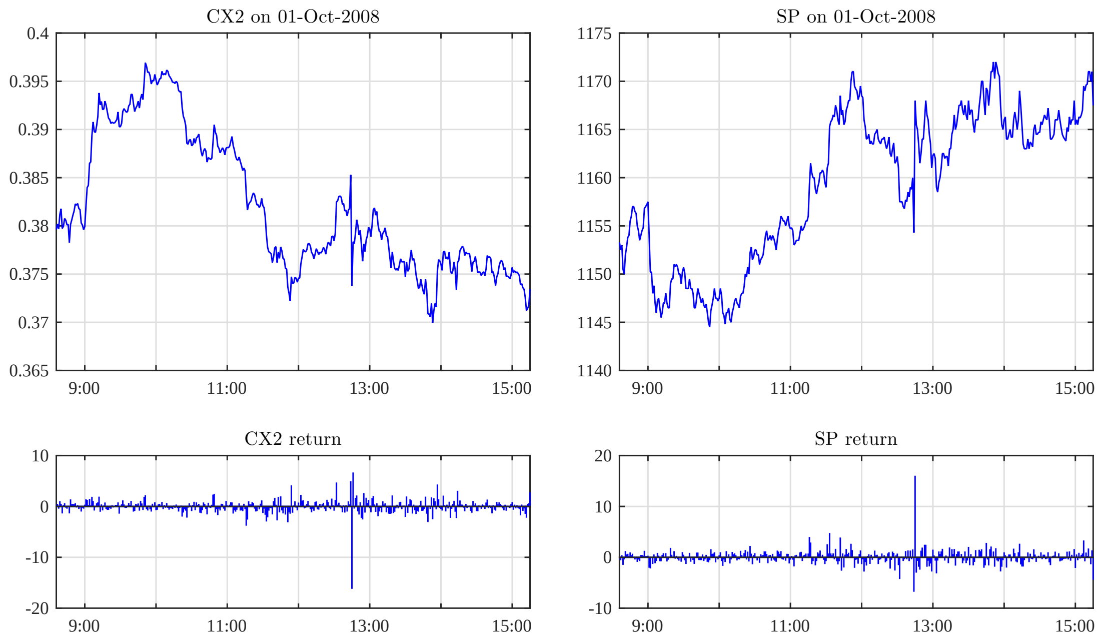
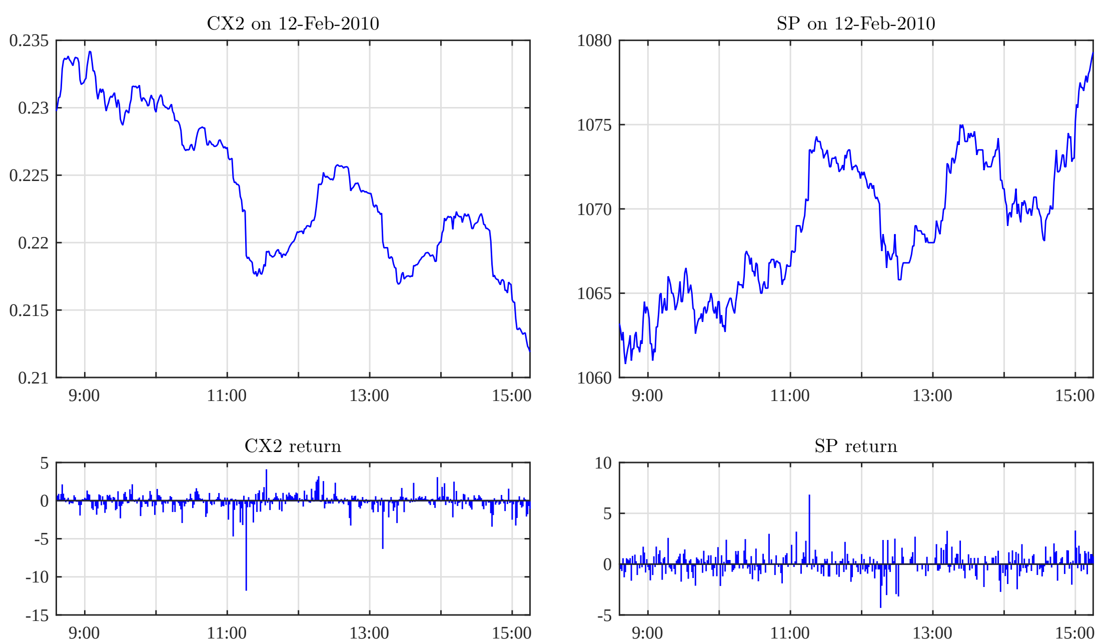



Figure 16: **Negative Jump of CX2 on February 12, 2010.** The top panels depict levels of CX2 and S&P 500, while the bottom panels depict the corresponding 1-minute returns, measured in multiples of sigma. At 11:15, CX2 has a big negative jump of -11.7 sigmas, while S&P 500 has a positive jump of 6.8 sigmas.




The conclusion that volatility can move sharply down (not just up) is supported by Figures 15 and 16, which illustrate two days in our sample where CX2 experiences a large *negative* jump. On October 1, 2008, CX2 has a negative jump of -16.1 sigmas. The jump appears to be *real*, as it is accompanied by a simultaneous *positive* jump of 15.9 sigmas in S&P 500 index. Similarly, on February 12, 2010, CX2 has a big negative jump of -11.7 sigmas, which is matched by a positive jump of 6.8 sigmas in S&P 500.

When artificial jumps in volatility indices are caused by idiosyncratic shifts in the effective range, they are likely to be negatively correlated. For example, a sudden narrowing of the strike range will often be followed by a return to the "normal" range. The support of this reversion hypothesis is provided by Table 7, which reports autocorrelation coefficients for two cases: when returns of volatility indices exceed 6 or 9 sigmas. In the former case, autocorrelation coefficient $\rho$ is negative and statistically highly significant for RX1, RX2, VIX$^*$, and VIX. When more extreme returns are considered in Panel B, the autocorrelation coefficient becomes even more negative. For VIX, it reaches -0.59, suggesting that after a big jump, the index experiences another big jump of the opposite sign. In contrast, $\rho$ is statistically zero for the CX indices.

## 6.2 Jump Autocorrelation

In this subsection, we investigate autocorrelation between large moves of volatility indices. When artificial jumps in volatility indices are caused by idiosyncratic shifts in the effective range, they are likely to be negatively correlated. For example, a sudden narrowing of the strike range will often be followed by a return to the "normal" range. The support of this reversion hypothesis is provided by Table 7. Panels A and B report autocorrelation coefficients for two cases: when returns of volatility indices exceed 6 or 9 sigmas. In the former case, the autocorrelation coefficient $\rho$ is negative and statistically highly significant for RX1, RX2, VIX$^*$, and VIX. However, when more extreme returns are considered in Panel B, the autocorrelation coefficient becomes even more negative. For VIX, it reaches -0.59, suggesting that after a big jump, the index often experiences another big jump of the opposite sign. In contrast, $\rho$ is statistically zero for the CX indices.

We next refine our analysis by considering a subset of consecutive jumps that occur on same day. The motivation is that the reversal of the strike range could potentially happen overnight, weakining the negative correlation between consecutive jumps. Panels C and D demonstrate that, indeed, the reversal effect becomes stronger and the autocorrelation coefficient $\rho$ becomes even more negative for RX1, RX2, VIX$^*$, and VIX. For example, for jumps exceeding 9 sigmas, the autocorrelation coefficient for VIX becomes -0.85. In contrast, for the CX indices $\rho$ remains statistically zero (smaller jumps in Panel C) or it cannot be computed due to lack of observations (larger jumps in Panel D).

Table 7: **Autocorrelation of Extreme Returns ("Jumps"), 1-min Frequency**. The autocorrelation coefficient $\rho$ is computed for returns exceeding 6 sigmas or 9 sigmas in absolute value. Unlike Panels A and B, Panels C and D compute autocorrelation only from consecutive jumps that occur on same day. Each panel reports the number of observations Nobs (pairs of consecutive jumps), the autocorrelation coeffiecient $\rho$, and its $p$-value. In Panel D, statistics for indices CX1-CX3 are not reported because Nobs is less than 5.

**Panel A: Jumps $\geq$ 6 sigmas**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* | VIX |
|---|---|---|---|---|---|---|---|
| Nobs | 702 | 520 | 340 | 300 | 278 | 727 | 840 |
| $\rho$ | -0.31 | -0.25 | 0.13 | -0.01 | 0.01 | -0.26 | -0.48 |
| $p$-value | 0.00 | 0.00 | 0.02 | 0.87 | 0.81 | 0.00 | 0.00 |

**Panel B: Jumps $\geq$ 9 sigmas**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* | VIX |
|---|---|---|---|---|---|---|---|
| Nobs | 215 | 118 | 37 | 34 | 34 | 222 | 329 |
| $\rho$ | -0.38 | -0.43 | -0.17 | -0.04 | -0.06 | -0.37 | -0.59 |
| $p$-value | 0.00 | 0.00 | 0.32 | 0.85 | 0.76 | 0.00 | 0.00 |

**Panel C: Jumps $\geq$ 6 sigmas on Same Days**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* | VIX |
|---|---|---|---|---|---|---|---|
| Nobs | 400 | 259 | 133 | 105 | 99 | 425 | 526 |
| $\rho$ | -0.48 | -0.43 | 0.10 | 0.03 | 0.03 | -0.41 | -0.63 |
| $p$-value | 0.00 | 0.00 | 0.27 | 0.79 | 0.77 | 0.00 | 0.00 |

**Panel D: Jumps $\geq$ 9 sigmas on Same Days**

| | RX1 | RX2 | CX1 | CX2 | CX3 | VIX* | VIX |
|---|---|---|---|---|---|---|---|
| Nobs | 95 | 46 | 2 | 0 | 4 | 96 | 183 |
| $\rho$ | -0.75 | -0.71 | | | | -0.67 | -0.85 |
| $p$-value | 0.00 | 0.00 | | | | 0.00 | 0.00 |

## 6.3 Volatility of Volatility

In this subsection, we investigate the realized variance (RV) of the volatility indices. Specifically, we compute daily RV from 1-minute returns and decompose it into the variance due to regular moves (non-jumps), small jumps, and large jumps, so that

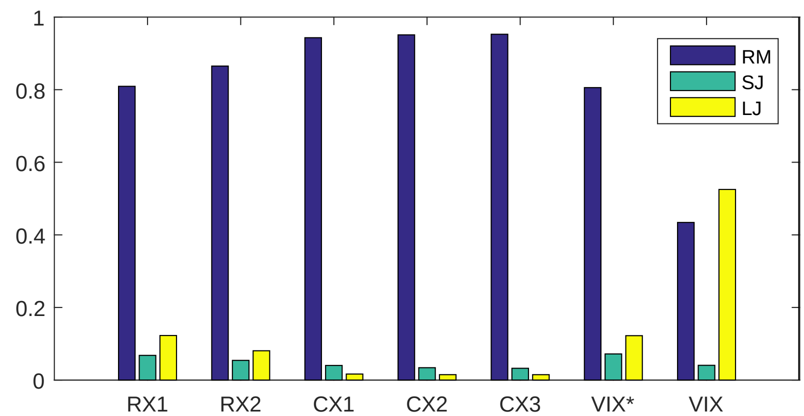
$$RV = RVRM + RVSJ + RVLJ.$$

Table 8 presents the results. The relative contributuion of large jumps to RV is considerable for indices for RX1, RX2, VIX*, and VIX. In particular, this contribution is 53% for VIX. In contrast, large jumps are rare for CX indices and their contribution to RV is only 1-2%. Figure 17 visualizes relative contributions to RV from regular moves, small jumps, and large jumps for different volatility indices.

Table 8: **Daily RV Decomposition** The Realized Variance is decomposed into the variance due to non-jumps, small jumps, and large jumps, $RV = RVRM + RVSJ + RVLJ$.

| | $RX1$ | $RX2$ | $CX1$ | $CX2$ | $CX3$ | $VIX^*$ | $VIX$ |
|---|---|---|---|---|---|---|---|
| RV: Mean | 0.46 | 0.41 | 0.44 | 0.47 | 0.48 | 0.46 | 0.85 |
| StdDev | 0.48 | 0.46 | 0.48 | 0.51 | 0.52 | 0.45 | 4.54 |
| Prop | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| RVRM: Mean | 0.38 | 0.36 | 0.42 | 0.44 | 0.45 | 0.37 | 0.37 |
| StdDev | 0.41 | 0.39 | 0.46 | 0.49 | 0.50 | 0.38 | 0.38 |
| Prop | 0.81 | 0.87 | 0.94 | 0.95 | 0.95 | 0.81 | 0.43 |
| RVSJ: Mean | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 |
| StdDev | 0.07 | 0.05 | 0.04 | 0.04 | 0.04 | 0.07 | 0.08 |
| Prop | 0.07 | 0.05 | 0.04 | 0.03 | 0.03 | 0.07 | 0.04 |
| RVLJ: Mean | 0.06 | 0.03 | 0.01 | 0.01 | 0.01 | 0.06 | 0.45 |
| StdDev | 0.20 | 0.20 | 0.06 | 0.06 | 0.07 | 0.20 | 4.43 |
| Prop | 0.12 | 0.08 | 0.02 | 0.01 | 0.01 | 0.12 | 0.53 |

Figure 17: **RV Decomposition.** Bars represent proportions of RM, SJ, and LJ contributions to Realized Variance.




# 7 Conclusions

We pursue two primary objectives. First, we seek to replicate the official real-time fifteen second VIX series from tick-by-tick quotes on S&P 500 option prices obtained from the official CBOE vendor. The purpose is to gauge the reliability of high-frequency movements in the VIX index and understand which features of the real-time series are robust to noise and measurement errors. This first step provides valuable insights into potential problems plaguing the VIX series. We identify a particularly important deficiency: the rule for determining the effective strike range used to compute the volatility index is not robust to common data errors or random shifts in market liquidity. This induces spurious jumps in VIX as well as any alternative index constructed on the basis of similar rules for the determination of the effective strike range. These "artificial" jumps have a profound impact on the high-frequency characteristics of the VIX returns. Importantly, the identical issue affects the daily closing values of the VIX which are used in numerous academic studies and are widely cited in the financial press. Day-to-day changes in the index are not reliable indicators of the corresponding *shifts* in implied volatility. On the other hand, all the implied volatility measures we explored convey near identical messages concerning the evolution in the overall *level* of volatility. Critical discrepancies arise only when focusing on the *changes* in the indices at an intraday or daily frequency.

Our second goal is to construct corridor implied volatility indices, CX, that retain the theoretical advantages of the model-free volatility concept while avoiding the pitfalls associated with the official VIX series. These CX indices may be computed directly from real-time option quotes via a simple and transparent computational rule. The main change from the VIX is the determination of a different truncation point for when to exclude far OTM options from the calculation. By ensuring intertemporal consistency in this dimension, we obtain CX indices that cover economically equivalent parts of the strike range. Consequently, they are internally consistent and the index values may be meaningfully compared across time, independently of the liquidity of the options market. As a result, we find that the number of spurious jumps in the volatility index are dramatically reduced, if not entirely eliminated, and the jump returns now consistently display a pronounced negative relation to the underlying equity returns. In short, the high-frequency CX and VIX return series display strikingly different time series characteristics along critical dimensions relevant for real-time financial decision-making and risk management. The same issues afflict the widely used daily VIX series, available from the CBOE website. The daily volatility returns implied by this VIX series are inherently noisy and caution should be exercised if the analysis is focused on the daily shifts in volatility. In such instances, our CX measures represent a superior alternative and they are readily computed from the underlying SPX option quotes.

Our findings point towards a number of issues for future research related to the use and construction of implied volatility measures. One important application is the identification and modeling of the equity variance risk premium. The premium is given by the wedge between the (expected) realized variance of the equity index returns and the MFIV measure over the corresponding horizon. In the literature, this is most often approximated by the difference between an (expected) realized variance measure, constructed from the cumulative sum of squared high-frequency equity-index returns, and the (squared) implied volatility index, represented by the CBOE VIX measure. Thus, to the extent VIX is a noisy, downward-biased, and excessively volatile measure of implied volatility, the quality of the risk premium measures are similarly degraded. It is evident that a precise and coherent identification of the variance risk premium requires the development of techniques to estimate the MFIV from options with strikes that span the full range of potential future values of the equity index. There are two obvious approaches to this issue. First, one may directly seek to extrapolate prices for deep OTM options so that the effective strike range is expanded to ensure all non-trivial contributions to the measure are included. Since observed quotes for far OTM options are unreliable, or even unobserved, such procedures hinge on specific assumptions about the shape and tail-fatness of the RND. One popular approach is to apply extrapolation in the Black-Scholes implied volatility space, either assuming implied volatility remains constant beyond the lowest and highest strike price for which reliable option values are available, as in Jiang & Tian (2005) and Carr & Wu (2009), or extending the implied volatility curves linearly from these extreme endpoints, as in Jiang & Tian (2007). A more recent sophisticated approach towards capturing the time-variation in the shape of the tails is provided by Bollerslev & Todorov (2011*a*), and it is also the subject of ongoing work in Andersen & Bondarenko (2011). Second, one may resort to exploring the variance risk premium only within ranges, or corridors, of the strike range for which we can reliably measure both the model-free implied variance and the corresponding realized return variation of the underlying equity-index returns. This approach provides a decomposition of the variance risk premium into segments of the return space, thus enabling direct identification of the size of the variance risk premium across different equity return states. For example, one may explore how much of the premium stems from high valuations of return variation in states where the equity prices are declining. This approach is explored in detail in **?**.

A second important avenue for future research is to enhance our understanding of the nature of the random variation in the quality and strike coverage of the SPX option quotes that constitute the most critical input to the computation of the implied volatility measures. The SPX options market is primarily a floor-based open outcry system. However, the electronic data feed providing the official SPX quotes does not originate from this market but are instead constructed from real-time quotes issued by two lead market maker firms at the CBOE as well as customer and broker orders listed on the same electronic platform. These electronically disseminated quotes generally have significantly wider spreads than the effective bid-ask spreads prevailing in the SPX trading pit. By studying the interaction of the quoting patterns from the lead market makers, customers and brokers, we may obtain a deeper understanding of the features in the data that induce noise as well as spurious jumps into the real-time VIX measures. As a consequence, improved procedures for extracting the relevant information from the real-time option quotes may be within reach. Gonzalez-Perez (2011) is currently exploring the relevant microstructure features of the SPX option market and the associated electronic quoting mechanism in order facilitate such direct inquiry into the origins of the distortions in the VIX index.

## References


Andersen, T.G. & O. Bondarenko. 2007. "Construction and Interpretation of Model-Free Implied Volatility." pp. 141–181. *Volatility as an Asset Class*, Edited by I. Nelken, London: Risk Books.

Andersen, T.G. & O. Bondarenko. 2011. "Estimating the Tails of the Risk-Neutral Distribution." *Working Paper in Progress, Northwestern University and University of Illinois at Chicago* .

Andersen, T.G., O. Bondarenko & M.T. Gonzalez-Perez. 2015. "Exploring Return Dynamics via Corridor Implied Volatility." *Review Financial Studies* 28:2902–2945.

Anderson, T.W. 1984. "The 1982 Wald Memorial Lectures: Estimating Linear Statistical Relationships." *Annals of Statistics* 12:1–45.

Bollerslev, T., G. Tauchen & H. Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review Financial Studies* 22:4463–4492.

Bollerslev, T. & V. Todorov. 2011*a*. "Estimation of Jump Tails." *Econometrica* 79:1727–1783.

Bollerslev, T. & V. Todorov. 2011*b*. "Tails, Fears and Risk Premia." *Journal of Finance* 66:2165–2211.

Bondarenko, O. 2000. "Recovering Risk-Neutral Densities: A New Nonparametric Approach." *Working Paper, University of Illinois, Chicago* .

Bondarenko, O. 2014. "Variance Trading and Market Price of Variance Risk." *Journal of Econometrics* 180:81–97.

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

Carr, P. & D.B. Madan. 1998. "Option Valuation Using the Fast Fourier Transform." *Journal of Computational Finance* 2:61–73.

Carr, P. & L. Wu. 2009. "Variance Risk Premiums." *Review Financial Studies* 22:1311–1341.

Dupire, B. 1993. "Model Art." *Risk* 6 (9):118–124.

Dupire, B. 1996. "A Unified Theory of Volatility." *Working Paper, BNP Paribus* .

Dupire, B. 2004. *"A Unified Theory of Volatility"*. *Derivatives Pricing: The Classic Collection*. (London: Risk Books). Ed. by P. Carr.

Gonzalez-Perez, M.T. 2011. "The Role of Market Makers in Determing the Quality of SPX Option Quotes." *Working Paper in Progress, Northwestern University* .

Jiang, G.J. & Y.S. Tian. 2005. "The Model-Free Implied Volatility and Its Information Content." *Review Financial Studies* 18:1305–1342.

Jiang, G.J. & Y.S. Tian. 2007. "Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index." *Journal of Derivatives* Spring:1–26.

Neuberger, A.J. 1994. "The Log Contract." *Journal of Portfolio Management* 20:74–80.

Todorov, V. 2010. "Variance Risk Premium Dynamics: The Role of Jumps." *Review Financial Studies* 23:345–383.

Todorov, V. & G. Tauchen. 2010. "Activity Signature Functions for High-Frequency Data Analysis." *Journal of Econometrics* 154:125–138.

# Appendix

## A Data and Data Filtering

### A.1 SPX Option Classes

We acquired the MDR (Market Data Retrieval) data for S&P 500 options from the CBOE subsidiary, Market Data Express (`http://www.marketdataexpress.com/`). The MDR data include tick-by-tick quotes and transactions throughout the trading day for all option classes issued by the CBOE on the S&P 500. Each option class is characterized by a letter code. We only consider options that the CBOE actually used in their computation of the VIX over our sample period, namely those in the SPB, SPQ, SPT, SPV, SPX, SPZ, SVP, SXB, SXM, SXY, SXZ, SYG, SYU, SYV and SZP categories. The latter are generally known as SPX equity options and they mature on the Saturday immediately following the third Friday of the expiration month (see `http://www.cboe.com/Products/EquityOptionSpecs.aspx`).

### A.2 Systemic Staleness in Option Quotes

By far, the most influential options for the computation of model-free implied volatility measures are the out-of-the-money (OTM) put and call options with strikes close to the at-the-money (ATM) strike which is given by the forward price. Hence, we actively monitor the liquidity of the ATM option as well as the set of twenty OTM put and twenty OTM call options closest to ATM for the two maturities exploited for the VIX computation. We label this group of options the "pivotal" ones.

Our systemic staleness filter flags episodes where there are no quote update among the pivotal options in the bid *or* the ask price at *either* of the two maturities for five minutes. Hence, this dummy variable equals one if all options within one of these four pivotal option groups have not been updated for five minutes and it is zero otherwise. When the flag is activated (equals unity), we classify the *entire* inactive period of five minutes or more as unreliable, and the volatility indices are not available (n.a.).

### A.3 Non-Convexity Filter

To preclude arbitrage opportunities, theoretical call and put prices must be monotonic and convex functions of the strike. In particular, the call prices must satisfy the following convexity restriction:

$$D_i \;=\; \frac{C(K_{i+1}) - C(K_i)}{K_{i+1} - K_i} - \frac{C(K_i) - C(K_{i-1})}{K_i - K_{i-1}} \;\geq\; 0\,,$$

and a similar restriction for the put prices. We obtain the option prices as the average of the bid and ask quotes and use the above restriction to identify "suspect" cross-section with apparent arbitrage violations, which could arise from recording errors, staleness, and other issues. Specifically, for each strike $K_i$, we compute the following measure of local non-convexity:

$$NC_i \;=\; -\min\{D_i, 0\}\,.$$

For low strikes ($K_i \leq F$), we compute $NC_i$ from OTM puts and, for high strikes ($K_i > F$), we use OTM calls. We then average $NC_i$ across all strikes to obtain the aggregate measure of non-convexity $NC$.

When $NC > 0.1$, we deem a cross-section unreliable and do not use it in our econometric analysis. For those cross-sections, the option prices indicate sizeable apparent arbitrage opportunities.

However, for some illustrations, we need to compute volatility indices even when the quality of data is poor ($NC > 0.1$). In those cases, prior to computing the volatility indices, we adjust option prices by running the so-called *Constrained Convex Regression* (CCR). This procedure has been implemented in Bondarenko (2000). Intuitively, CCR searches for the smallest (in the sense of least squares) perturbation of option prices that restores the no-arbitrage restrictions.

### A.4 Bounce-Back Filter

We detect situations where a volatility index experiences a large move, which is almost immediately offset by a jump of a similar magnitude, but in the opposite direction. Such behavior is usually indicative of a data error which impacts only a few recorded quotes. Such an error induces a jump in returns when it occurs, and then another one when the error is eliminated and the recorded quotes reverse to their appropriate level. Specifically, we label such an incident a " bounce-back" if the index jumps by at least 9 "sigmas" and either (i) more than 75% of this jump is reversed within one minute, or (ii) more than 80% of the jump is reversed within two minutes.

## B Robust Volatility Estimation

For a given volatility index, let the associated one-minute log return at time $t$ on trading day $d$ by $r_{dt}$. We assume this return may be represented as

$$r_{dt} \;=\; \sigma_d\, f_t\, z_{dt}\,,$$

where $\sigma_d$ is average volatility for trading day $d$, $f_t$ is a scaling factor which adjusts for the intraday volatility pattern, and $z_{dt}$ are i.i.d. random variables with zero mean and unit standard deviation, but not necessary normally distributed.

We estimate the daily volatilities $\{\sigma_d\}_1^{N_d}$ and intraday adjustment factors $\{f_t\}_1^{N_t}$ as follows. First, for fixed $t$, we compute the average squared return across all trading days,

$$\frac{1}{N_d}\sum_{d}^{N_d} r_{dt}^2 \;=\; f_t^2 \cdot \frac{1}{N_d}\sum_{d}^{N_d} \sigma_d^2\, z_{dt}^2\,.$$

Since $z_{dt}$ are i.i.d. with $E[z_{dt}] = 0$ and $\mathrm{Var}(z_{dt}) = 1$, we obtain,

$$E\left[\frac{1}{N_d}\sum_{d}^{N_d} r_{dt}^2\right] \;=\; f_t^2 \cdot \left(\frac{1}{N_d}\sum_{d}^{N_d} \sigma_d^2\right) \;=\; V_{mean} \cdot f_t^2\,,$$

where $V_{mean}$ is a constant independent of $t$. Therefore, the adjustment factor $f_t$ may be estimated as

$$f_t^2 \;=\; \frac{\frac{1}{N_d}\sum_d^{N_d} r_{dt}^2}{V_{mean}}\,,$$

where $V_{mean}$ now is determined from the condition that,

$$\sigma_d^2 \;=\; \frac{1}{N_t} E\left[\sum_t^{N_t} r_{dt}^2\right] \;=\; \frac{1}{N_t}\sum_t^{N_t} \sigma_d^2 \cdot f_t^2\,,$$

or

$$\frac{1}{N_t}\sum_t^{N_t} f_t^2 \;=\; 1\,. \tag{8}$$

The above approach provides an unbiased estimator of $f_t^2$. However, it might not be robust in the presence of extreme returns in the $r_{dt}$ series. Therefore, we implement a robustified version of the same approach, based on the sample median rather than the mean. Specifically,

$$E\left[\text{Median}_d\left\{r_{dt}^2\right\}\right] \;=\; f_t^2\,\text{Median}_d\left\{\sigma_d^2\right\} \;=\; V_{med}\cdot f_t^2\,,$$

where $V_{med}$ again denotes a constant independent of $t$. The adjustment factor $f_t$ i then estimated as

$$f_t^2 \;=\; \frac{\text{Median}_d\left\{r_{dt}^2\right\}}{V_{med}},$$

subject to the constraint in (8). For additional robustness, the estimated intraday factors are smoothed by averaging them over 10-minute windows.[21] Armed with $\{f_t\}_1^{N_t}$, the daily volatility $\sigma_d$ can now be estimated as the volatility of the re-scaled returns, $u_{dt} \;=\; \frac{r_{dt}}{f_t}$. This could be done in many ways, but we focus on the robust estimator based on 5-to-95 percentile range. Specifically, we sort the re-scaled returns $u_{dt}$ for a given day $d$ and determine their 5- and 95-percentiles, $P(0.05)$ and $P(0.95)$. Then,

$$\sigma_d \;=\; \frac{P(0.95)-P(0.05)}{3.2898},$$

where the denominator equals the 5-to-95 percentile range of the standard normal random variable.[22]

Finally, when defining large index returns, we take into account both the daily measure of volatility $\sigma_d$ and the intraday adjustment factor $f_t$. That is, the move is deemed large if the absolute value of the ratio,

$$\left|\frac{r_{dt}}{\sigma_d\, f_t}\right| \;=\; \left|\frac{u_{dt}}{\sigma_d}\right|$$

exceeds a pre-specified threshold.

# C Orthogonal Regression

The Orthogonal regression (OR), also known as Total Least Squares (TLS) regression, has a long history in statistics and economics, see, e.g. the discussion in Anderson (1984). It has been viewed as more appropriate than the Ordinary Least Squares (OLS) regression in some circumstances, including when both the predictor and response variables are measured with error. In the bivariate case, OR minimizes the sum of squared perpendicular (or total) distances from the data to the fitted regression line, in contrast to OLS which minimizes the sum of squared vertical distances. The slope of OR line is related to the first Principal Component.

In Figure **??** we regress volatility index returns onto the underlying futures return, using OR. When a volatility index contains large "artificial" jumps, those observations tend to be located along the vertical axis (a small move

[21]For all volatility indices, the intraday factors range from about 0.6 to 2.0.

[22]We also experimented with other robust estimators of $\sigma_d$, and the results are quantitatively similar.

in the underlying futures but a very large return in the volatility index). The slope of OLS regression has a very low sensitivity to such outliers and is less informative for our purposes. In contrast, the slope of OR is affected by the presence of extreme volatility returns along the vertical axis.