Truncation error stabilization for model-free implied moment estimator*

Geul Lee and Li Yang†

School of Banking and Finance, University of New South Wales

3 June 2016

Abstract

This study suggests a new method to mitigate the impact of truncation, i.e., the unavailability of extremely deep-out-of-the-money option quotes, on the model-free implied moment estimator of Bakshi et al. (2003). The estimation error due to truncation, i.e., truncation error, can be troublesome when comparing the implied moments cross-sectionally or tracking the time-series dynamics of the moments especially when the error is volatile so that the differences between true moments are less visible. We show how the methodology of Dennis and Mayhew (2002) can be improved to reduce the impact of truncation error by making the error less volatile.

*We thank Tom Barkley, Guanglian Hu, Jungmu Kim, Jaehoon Lee, Yang Liu, Ji Yeol Jimmy Oh, Thijs van der Heijden, and Yingying Wu for helpful comments and suggestions, as well as the seminar participants at 2014 Australasian Finance and Banking Conference, 2015 Financial Markets and Corporate Governance Conference, 2015 FMA Asian Conference, 2015 International Finance and Banking Society Conference, 2015 Asian Finance Association Annual Meeting, 2015 Derivatives Markets Conference, 2015 FMA Annual Meeting, and 2015 Conference on Asia-Pacific Financial Markets. An earlier version of the paper was circulated under the title "Impact of truncation on model-free implied moment estimator". This research includes computations using the Linux computational cluster Katana supported by the Faculty of Science, University of New South Wales.

†Corresponding authors. E-mail addresses: geul.lee@unsw.edu.au and l.yang@unsw.edu.au

Truncation error stabilization for model-free implied moment estimator

Abstract

This study suggests a new method to mitigate the impact of truncation, i.e., the unavailability of extremely deep-out-of-the-money option quotes, on the model-free implied moment estimator of Bakshi et al. (2003). The estimation error due to truncation, i.e., truncation error, can be troublesome when comparing the implied moments cross-sectionally or tracking the time-series dynamics of the moments especially when the error is volatile so that the differences between true moments are less visible. We show how the methodology of Dennis and Mayhew (2002) can be improved to reduce the impact of truncation error by making the error less volatile.

JEL Classification: C14; C58; G13.

Keywords: Truncation error; model-free implied moment estimators; domain stabilization.

1 Introduction

The implied moment estimators of Bakshi et al. (2003) are in active use to capture the characteristics of the implied risk-neutral density (RND), especially by recent studies in the areas of asset pricing and portfolio allocations. Since the implied moments reflect market participants' expectations of underlying asset return, the estimators are adopted in order to measure investor sentiment (Han, 2008), improve portfolio selection (DeMiguel et al., 2013), explain the cross-section of expected asset returns (Conrad et al., 2013).

While the estimators of Bakshi et al. (2003) retain theoretical attractiveness because they do not rely on any assumptions on the dynamics of underlying price process, the limited availability of out-of-the-money (OTM) option quotes causes an empirical issue when these estimators are employed. To estimate the implied moments using the estimators of Bakshi et al. (2003), one requires OTM option prices for the continuum of strike prices from zero to positive infinity. In reality, however, option quotes are observable only for a discrete and finite set of strike prices. Jiang and Tian (2005) point out two issues that are induced by this limited availability of option quotes. First, since option quotes are available only for a discrete set of strike prices, the integrals of weighted option prices that are used for the implied moment estimation need to be approximated. Second, since option quotes are completely unavailable for the extremely low or high strike prices, the integration of weighted option prices cannot be conducted for the strike price domain from zero to positive infinity. While the estimation error due to the strike price discreteness can be mitigated relatively easily by employing interpolation techniques, it is more difficult to deal with the estimation error due to the complete unavailability of deep-in-the-money (DITM) and deep-out-of-the-money (DOTM) option quotes. Jiang and Tian (2005) name the second type of estimation errors "truncation errors" since it looks as if some options quotes have been "truncated".

So far, two truncation error reduction methods have been suggested by previous literature. First, Jiang and Tian (2005) propose linear extrapolation (LE), which is conducted by extending the Black-Scholes implied volatility curve by assuming that the implied volatility is constant beyond the minimum and maximum strike prices, and then generate option prices whose prices correspond to the extended curve. Second, Dennis and Mayhew (2002) suggest domain symmetrization (DSym), which is employed by additionally discarding option prices to make the minimum and maximum strike prices equidistant from the underlying price. Dennis and Mayhew (2002) argue that DSym makes the implied skewness estimator less biased since the fair value estimate of the moment-related contracts, of which the implied moment estimators of Bakshi et al. (2003) are nonlinear functions, depends on the difference between the weighted average prices of OTM calls and puts.

Since Lee (2015) examines LE in detail and shows that there is a need to further develop the methodology to address truncation error, we focus on DSym and investigate whether this method is indeed effective, as well as whether it can be improved. We start from the question of whether DSym remains effective even when the implied RND is asymmetric. Dennis and Mayhew (2002) show that DSym can reduce the truncation error of the implied skewness estimator of Bakshi et al. (2003) while assuming that the underlying price follows a constant volatility process with no jumps, with which the implied RND is symmetric. Several empirical studies, however, show that the implied RND is asymmetric in most options markets. Hence, we examine the effectiveness of DSym under the circumstances where the implied RND is asymmetric. This study reveals that

DSym can be misleading when the implied RND is asymmetric.

Next, we suggest how DSym can be improved to reduce the impact of truncation even when the implied RND is asymmetric. Our main idea is to change the objective from 'minimizing the error' to 'stabilizing the error'. Dennis and Mayhew (2002) argue that if the size of estimation bias is stable across all observations, one can discern differences among the observations even when it is relatively difficult to estimate the true value. Hence, if finding out the true value of the implied moment estimation is not the major concern, and the estimation is conducted mainly for cross-sectional or time-series comparison, making the truncation error less volatile can be more effective than merely minimizing the mean of the error without concerning the volatility of the error. Given this idea, we first investigate the relationship between the level of truncation and the magnitude of truncation error, and then show that one can stabilize the latter by controlling the former based on the relationship. We call this new method domain stabilization (DStab).

The rest of this study is constructed as follows. Section 2 summarizes the definition of option-implied moment estimators in Bakshi et al. (2003). Section 3 describes the S&P 500 index options data used in this study. In Section 4, we report the results of our analysis based on generated option prices. Section 5 summarizes the empirical results. Section 6 introduces DStab and tests its effectiveness. Section 7 discusses the main findings and conclusions.

2 Model-free option-implied moment estimators

To estimate the variance, skewness, and kurtosis of the implied RND, Bakshi et al. (2003) first define three moment-related contracts, i.e., volatility contract V, cubic contract W, and quartic contract X, whose payoffs at maturity are equivalent to the second, third, and fourth-powers of holding period log-return, respectively. Bakshi et al. (2003) demonstrate that the fair value of three contracts under the risk-neutral measure can be defined using OTM option prices as

$$\begin{aligned} V(t,\tau) = &\int_{S(t)}^{\infty} \frac{2\left(1-\ln\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau;K)dK \\ &+ \int_{0}^{S(t)} \frac{2\left(1+\ln\left[\frac{S(t)}{K}\right]\right)}{K^2} P(t,\tau;K)dK, \end{aligned} \tag{1}$$

$$W(t,\tau) = \int_{S(t)}^{\infty} \frac{6\ln\left[\frac{K}{S(t)}\right] - 3\left(\ln\left[\frac{K}{S(t)}\right]\right)^2}{K^2} C(t,\tau;K)dK$$
$$- \int_{0}^{S(t)} \frac{6\ln\left[\frac{S(t)}{K}\right] + 3\left(\ln\left[\frac{S(t)}{K}\right]\right)^2}{K^2} P(t,\tau;K)dK, \tag{2}$$

$$X(t,\tau) = \int_{S(t)}^{\infty} \frac{12\left(\ln\left[\frac{K}{S(t)}\right]\right)^2 - 4\left(\ln\left[\frac{K}{S(t)}\right]\right)^3}{K^2} C(t,\tau;K)dK$$
$$+ \int_{0}^{S(t)} \frac{12\left(\ln\left[\frac{S(t)}{K}\right]\right)^2 + 4\left(\ln\left[\frac{S(t)}{K}\right]\right)^3}{K^2} P(t,\tau;K)dK, \tag{3}$$

where $C(t,\tau;K)$ and $P(t,\tau;K)$ denote the call and put prices, for strike price K, time t, and maturity τ, respectively. Then the volatility, skewness, and kurtosis of the implied RND for time t and maturity τ can be derived as

$$\mathrm{VOL}(t,\tau) = \left[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right]^{1/2}, \tag{4}$$
$$\mathrm{SKEW}(t,\tau) = \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2]^{3/2}}, \tag{5}$$
$$\mathrm{KURT}(t,\tau) = \frac{e^{r\tau}X(t,\tau) - 4\mu(t,\tau)e^{r\tau}W(t,\tau) + 6e^{r\tau}\mu(t,\tau)^2V(t,\tau) - 3\mu(t,\tau)^4}{[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2]^2}, \tag{6}$$

where

$$\mu(t,\tau) \equiv E_t^* \ln\left[\frac{S(t+\tau)}{S(t)}\right] = e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau). \tag{7}$$

3 Data

The S&P 500 index options data used in this study span the eleven-year sample period from January 2000 to December 2010. The daily option quotes and risk-free yield curve data are retrieved from the IvyDB OptionMetrics database via Wharton Research Data Services (WRDS). The closing option price is estimated as the mid-point between the closing bid and ask prices. The risk-free rate for each day and maturity is estimated by linearly interpolating the two nearest points on the daily yield curve. OptionMetrics also provides the option-implied dividend rate for each trading day, and this rate is used to approximate the dividend rate $q(t,T)$ for day t_0 and maturity date T

as

$$q(t_0, T) = \left[\prod_{i=1}^{n}(1 + q^*(t_i))\right]^{1/n} - 1, \tag{8}$$

where n is the number of implied dividend rates available for days between t_0 and T, $q^*(t_i)$ is the implied dividend rate for day t_i, which is between t_0 and T.

The following filtration conditions are applied to remove inadequate observations: (1) observations with any missing data entry are removed; (2) observations are excluded if the maturity is shorter than one week or longer than one year; (3) observations are included only if the sum of daily trading volume for the corresponding maturity date is non-zero; (4) observations are removed if the bid price is zero or higher than the ask price; (5) observations that violate the no-arbitrage restrictionare excluded; (6) observations with the mid-point price less than 0.375 are removed; and (7) observations with the bid-ask spread larger than the mid-point price are excluded. Table 2 summarizes the final sample properties.

4 Investigation with generated option prices

In this section, we investigate the impact of truncation with the option prices that are generated using option pricing models. Section 4.1 explains how the option prices are generated. Section 4.2 evaluates the effectiveness of DSym.

4.1 Generation of option prices

We generate two sets of option prices using the Black-Scholes constant volatility (BS) model and the stochastic volatility and jump (SVJ) model of Bakshi et al. (1997), respectively. BS model is chosen to set a benchmark as well as to link this study to Dennis and Mayhew (2002) and Jiang and Tian (2005) both of which also employ BS model. On the other hand, SVJ model is adopted to generate a more realistic simulation setting as in Jiang and Tian (2005).

For SVJ model, the underlying price is assumed to follow a process

$$\frac{dS(t)}{S(t)} = [r - \lambda\mu_J]dt + \sqrt{V(t)}d\omega_S(t) + J(t)dq(t), \tag{9}$$

$$dV(t) = [\theta_v - \kappa_v V(t)]dt + \sigma_v \sqrt{V(t)}d\omega_v(t), \tag{10}$$

$$\ln[1+J(t)] \sim N(\ln[1+\mu_J] - 0.5\sigma_J^2, \sigma_J^2), \tag{11}$$

where r is the constant risk-free interest rate, λ is the frequency of jumps per year, $V(t)$ is the part of return variance that is due to diffusion process, $\omega_S(t)$ and $\omega_v(t)$ are standard Brownian motions with $\mathrm{Cov}[d\omega_S(t), d\omega_v(t)] = \rho dt$, $J(t)$ is the percentage jump size that is log-normally i.i.d. over time as in (11), μ_J is the unconditional mean of $J(t)$, $\sigma_J(t)$ is the standard deviation of $\ln[1+J(t)]$, $q(t)$ is a Poisson jump counter with intensity λ so that $P[dq(t)=1] = \lambda dt$ and $P[dq(t)=0] = (1-\lambda)dt$, κ_v is the speed of adjustment of $V(t)$, θ_v/κ_v is the long-term mean of $V(t)$, and σ_v is the volatility of $V(t)$.

To generate option prices, we first need to set the value of model parameters. To maintain the analysis as realistic as possible, we set the parameter values based on the daily model calibration results on the S&P 500 index options data. Following Bakshi et al. (1997), if there are n call prices and m put prices that are observed on day t, we find the parameter vector $\boldsymbol{\phi}(t)$ for each model which solves

$$\min_{\boldsymbol{\phi}(t)} \left[\sum_{i=1}^{n} |C^*(t,\tau_i;K_i) - C(t,\tau_i;K_i)|^2 + \sum_{j=1}^{m} |P^*(t,\tau_j;K_j) - P(t,\tau_j;K_j)|^2 \right], \tag{12}$$

where $C^*(t,\tau_i;K_i)$ and $C(t,\tau_i;K_i)$ are the observed and model prices of ith call option with time to maturity τ_i and strike price K_i, and $P^*(t,\tau_j;K_j)$ and $P(t,\tau_j;K_j)$ are the observed and model prices of jth put option with time to maturity τ_j and strike price K_j, respectively.[1] After having the daily calibration results, we then calculate the mean of daily parameter vectors for the entire T trading days in our sample period, which can be defined as

$$\boldsymbol{\Phi} \equiv \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\phi}(t). \tag{13}$$

This mean vector $\boldsymbol{\Phi}$ for each model is employed as the parameter vector to generate simulated option prices. Next, we calculate model price of OTM options while fixing the maturity to be three months. The maturity of three months is chosen to represent the maturities of two and four months that are considered in Section 5. The elements in $\boldsymbol{\Phi}$ are listed in the first column of Table 1

[1] See the appendix of Bakshi et al. (1997) to find the characteristic functions that constitute a closed-form solution for model call price in SVJ case. In this study, model price for puts are calculated using the put-call parity relationship.

with relevant summary statistics. The table shows that SVJ model can explain the option price data much better, which is consistent with Bakshi et al. (1997) and several other related studies. Figure 1 illustrates some basic properties of the generated option prices.

4.2 Domain symmetry and implied skewness estimator

This subsection investigates the effectiveness of the DSym. We first show that the integration domain symmetry should be defined in terms of log-moneyness rather than strike price in Section 4.2.1. Based on this finding, we examine the relationship between the level of domain asymmetry and the implied skewness estimate in Section 4.2.2.

4.2.1 Truncation in terms of log-moneyness

If OTM option prices are available only for a strike price domain $[K_{\min}(t,\tau), K_{\max}(t,\tau)]$ for time t and maturity τ, where $0 \leq K_{\min}(t,\tau) \leq S(t) \leq K_{\max}(t,\tau) < \infty$, and integrations are conducted only for this domain for the fair value estimation of V, W, and X, then it is equivalent to assuming that the OTM option price is zero for the strike price domains $[0, K_{\min}(t,\tau))$ and $(K_{\max}(t,\tau), \infty)$, i.e., $P(t,\tau;K) \equiv 0$ for $\{K : 0 < K < K_{\min}(t,\tau)\}$ and $C(t,\tau;K) \equiv 0$ for $\{K : K_{\max}(t,\tau) < K < \infty\}$. The following proposition shows how this assumption affects the fair value estimation:

Proposition 1. If truncation exists for the strike price domain $(0, K_{\min}(t,\tau))$ and $(K_{\max}(t,\tau), \infty)$, and the fair value of V, W, and X are estimated without considering the OTM option prices on the truncated domain, it is equivalent to assuming that, given the no-arbitrage condition, for the risk-neutral probability measure $\mathbb{P}^*$,

$$\mathbb{P}_t^* \left\{ \ln\left[\frac{S(t+\tau)}{S(t)}\right] < \ln\left[\frac{K_{\min}(t,\tau)}{S(t)}\right] \right\} = 0; \quad \text{and} \quad \mathbb{P}_t^* \left\{ \ln\left[\frac{S(t+\tau)}{S(t)}\right] > \ln\left[\frac{K_{\max}(t,\tau)}{S(t)}\right] \right\} = 0, \tag{14}$$

where $\mathbb{P}_t^*\{\cdot\}$ is the conditional probability operator for the measure $\mathbb{P}^*$ with respect to the filtration $\mathcal{F}_t$ for time t.

Proof. Since the option payoff is always non-negative, there will be an arbitrage opportunity if a call price $C(t,\tau;K)$ for time t, maturity τ, and strike price K is zero but $\mathbb{P}(S_{t+\tau} - K > 0) > 0$. Similarly, there will be an arbitrage opportunity if a put price $P(t,\tau;K)$ is zero but $\mathbb{P}(K - S_{t+\tau} > 0) > 0$.

Hence, assuming both no-arbitrage condition and zero option price is equivalent to assuming that the probability that the option will become profitable is zero. □

In this study, Proposition 1 is assumed to cause a truncation of the implied RND, given that some of the unavailable DOTM option prices are in fact observed as nonzero value but discarded during the data filtration process in almost every case. This means that at least for some of the DOTM options for which the option price is regarded as unavailable, market participants think that it is possible for the underlying price to reach the corresponding strike prices at maturity. In addition, based on Proposition (1) which shows that the locations at which the implied RND is truncated due to the option price unavailability are equal to the log-moneyness of endpoint strike prices, i.e., $\ln\left(K_{\min}(t,\tau)/S(t)\right)$ and $\ln\left(K_{\max}(t,\tau)/S(t)\right)$, we introduce a log-moneyness-based definition of integration domain symmetry in Section 4.2.2.

4.2.2 Domain asymmetry and implied skewness estimate

To examine the relationship between domain asymemtry and implied skewness estimate, we suggest an alternative definition of domain symmetry that is based on log-moneyness of endpoint strike prices, based on the findings in Section 4.2.1. in contrast to Dennis and Mayhew (2002) who define the integration domain symmetry as a state where the minimum strike price $K_{\min}$, underlying price S, and maximum strike price $K_{\max}$ satisfy a condition $S - K_{\min} = K_{\max} - S$, the log-moneyness-based integration domain symmetry is defined as the state where a condition $\ln(S/K_{\min}) = \ln(K_{\max}/S)$ does hold.

Based on the two definitions, we estimate the implied skewness while setting the level of domain asymmetry differently using an asymmetry coefficient. When the integration domain symmetry is defined in terms of strike price, for a domain half-width W which satisfies $K_{\max} - K_{\min} = 2W$, and an asymmetry coefficient c, we control the level of integration domain asymmetry by setting $K_{\min}$ and $K_{\max}$ as $S-(1-c)W$ and $S+(1+c)W$, respectively. Under this specification, the ratio of the distances from the underlying price to the minimum and maximum strike prices in unit of strike price becomes $S - K_{\min} : K_{\max} - S = 1 - c : 1 + c$. On the other hand, when the integration domain symmetry is defined in terms of log-moneyness, we set $K_{\min}$ and $K_{\max}$ in a way that the following conditions hold:

$$(1+c)\ln(S/K_{\min}) = (1-c)\ln(K_{\max}/S); \quad \text{and} \quad K_{\max} - K_{\min} = 2W.$$

Here the ratio of the distances from the underlying price to the minimum and maximum strike prices in unit of log-moneyness becomes $\ln(S/K_{\min}) : \ln(K_{\max}/S) = 1-c : 1+c$. In both cases, the integration domain is regarded as symmetric when $c = 0$, biased to the OTM put side when $c < 0$, and biased to the OTM call side when $c > 0$.

Figure 2 demonstrates the estimation results. Two interesting points can be found. First, when BS model is considered so that the implied RND is symmetric, the truncation error of the implied skewness estimator is shown to be larger when the integration domain is more asymmetric, which is consistent with Dennis and Mayhew (2002). However, in Figure 2a in which the symmetry is defined in terms of strike price difference, it can be found that the estimate converges to the true value most quickly when the domain is slightly biased to the OTM call side. In contrast, in Figure 2b in which the symmetry is defined in terms of log-moneyness, the estimate is shown to converge to the true value most quickly when the domain is symmetric, which is consistent with Proposition 1. Second, in the SVJ cases where the implied skewness is non-zero, the truncation error can be smaller when domain is more asymmetric. In Figures 2c and 2d in which the implied skewness is negative, it can be found that the truncation error tends to be smaller when the domain is more biased to the OTM put side. In contrast, in Figures 2e and 2f for which the sign of the mean jump size μ and correlation coefficient ρ in SVJ model are switched to be positive so that implied skewness is also positive, it is shown that the truncation error tends to decrease as the domain becomes more biased to the OTM call side.

The results in this subsection have two implications. First, although DSym minimizes the truncation error when the implied skewness is zero and the integration domain symmetry is defined in terms of log-moneyness, it may not if one of these conditions does not hold. This also implies that the effectiveness of DSym depends on the true level of implied skewness. Second, given that the width of integration domain can decrease significantly by DSym, especially when the raw domain is severely asymmetric, truncation error can be decreased by allowing the integration domain to be asymmetric and preserving the width of integration domain when the implied skewness is supposed to be nonzero.

5 Empirical analysis

In this section, we conduct a set of empirical analyses on the S&P 500 index options data to further discuss the findings in the previous sections. Section 5.1 describes how option prices are estimated for fixed maturities by generating implied volatility surface. Section 5.2 explains how the size of truncation error is estimated from a proxy variable. Section 5.3 analyzes the relationship between the width of integration domain and the truncation error size. Section 5.4 investigates the relationship between the asymmetry level of integration domain and the truncation error size.

5.1 Generation of implied volatility surface

Following Jiang and Tian (2005), we only consider a number of fixed maturities by extracting implied volatility curves from daily implied volatility surface, in order to mitigate the telescoping problem that is pointed out by Christensen et al. (2002).[2] To generate a surface, we first collect the Black-Scholes implied volatilities from all observable OTM option prices. Next, we estimate a bicubic spline function with the use of the implied volatility observations. Finally, we generate a truncated implied volatility curve for each maturity using the bicubic spline function and the minimum and maximum strike prices. If the minimum and maximum strike prices are not observable for a maturity, they are approximated by linearly interpolating the corresponding endpoint strike prices for the two nearest maturities for which option prices are observable. Finally, the implied volatility curve is converted to a set of OTM option prices with the strike price interval of 0.1. If LE is employed, the implied volatility level at the minimum and maximum strike prices are extrapolated up to the points where the strike prices are $S(t)/3$ and $3S(t)$, respectively, where $S(t)$ is the dividend-adjusted index level on day t. The strike price interval between the OTM option prices that are generated by LE is also set to 0.1.

[2] Christensen et al. (2002) point out that given the fixed maturity dates, time to maturity for an option is telescoping, i.e., decreasing over time, and therefore makes the time period between present date and maturity date overlapping for option samples in different days but with the same maturity date.

5.2 Estimating the size of truncation error

Since the true value of implied moments is unknown for the option prices that are observed from markets, a proxy variable needs to be employed to approximate the size of truncation error. We adopt the absolute percentage change (APC) in implied moment estimate after LE as a proxy for truncation error. As shown in Section 4, LE is shown to reduce the size of truncation error significantly, although not completely. Hence, it can be conjectured that the absolute percentage change in moment estimate after extrapolation increases as truncation error in percentage increases. We define APC as

$$\text{APC} = \left| \frac{(\text{Estimate with LE}) - (\text{Estimate without LE})}{(\text{Estimate without LE})} \right|, \tag{15}$$

and use this variable to approximate the size of truncation error for the rest of this section. Table 3 reports some preliminary statistics of APC for the implied volatility, skewness, and kurtosis estimators. Table 3 shows that the mean of APC for the implied kurtosis estimator is much larger than the one for the implied skewness estimator, and even larger than the one for the implied volatility estimator. This is consistent with the result in Section 4 that the implied skewness and kurtosis estimators are in fact affected by truncation more significantly than the implied volatility estimator is.

5.3 Width of integration domain and truncation error size

In this subsection, we explore the relationship between strike price domain width and magnitude of truncation error. Section 4.2.1 suggests that the endpoint log-moneyness of the truncated integration domain is closely related to the endpoint of the truncated implied RND. In addition, Section 4.2 shows that the level of implied volatility affects the relationship. Based on these findings, we define the width of integration domain in a number of different ways, and investigate how closely the size of truncation error is related to the width of integration domain when the width is defined in different terms.

We introduce four definitions of integration domain width as follows:

$$(\text{Strike price width}) = K_{\max}(t,\tau) - K_{\min}(t,\tau), \tag{16}$$

$$(\text{Moneyness width}) = \frac{K_{\max}(t,\tau)}{S(t)} - \frac{K_{\min}(t,\tau)}{S(t)}, \tag{17}$$

$$(\text{Log-moneyness width}) = \ln\left(\frac{K_{\max}(t,\tau)}{S(t)}\right) - \ln\left(\frac{K_{\min}(t,\tau)}{S(t)}\right), \tag{18}$$

$$(\text{Volatility-adjusted log-moneyness width}) = \frac{(\text{Log-moneyness width})}{\sqrt{\tau}\cdot \mathrm{VOL}(t,\tau)}, \tag{19}$$

where $\mathrm{VOL}(t,\tau)$ is the implied volatility estimate after LE. Roughly speaking, the last measure gauges the width of domain in unit of implied standard deviation of log-return.

In order to examine the relationship between domain width and truncation error size, we conduct a set of univariate linear regression analyses where each of the measures above is employed as independent variables. APC is employed as the dependent variable. Table 4 reports the regression result for the two different maturities, i.e., two and four months. As expected, the table suggests a negative relationship between the integration domain width and the truncation error size. A notable point is that the width of integration domain is shown to explain the size of truncation error much better when the width is defined in terms of volatility-adjusted log-moneyness.[3] Furthermore, it is shown in Column [5] that this high explanatory power cannot be fully obtained when only implied holding period volatility is adopted as an independent variable. This implies that it is important to consider both the log-moneyness of endpoint strike prices and the level of implied volatility when explaining the relationship between the width of integration domain and the size of truncation error.

5.4 Asymmetry level of integration domain and truncation error size

Section 4.2 shows that integration domain symmetry does not always lead to smaller truncation error, especially when the true implied skewness is far from zero. This section verifies this empirically by investigating the relationship between the asymmetry level of integration domain and the size of truncation error.

[3] In an unreported analysis, it is found that adjusted R^2 becomes smaller if we replace volatility-adjusted log-moneyness with volatility-adjusted strike price or volatility-adjusted moneyness.

As in Section 5.3, we employ a number of different definitions of the asymmetry level:

$$\text{(Strike price width difference)} = \frac{(K_{\max}(t,\tau) - S(t)) - (S(t) - K_{\min}(t,\tau))}{1000 \cdot \sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \tag{20}$$

$$\text{(Moneyness width difference)} = \frac{(K_{\max}(t,\tau)/S(t) - 1) - (1 - K_{\min}(t,\tau)/S(t))}{\sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \tag{21}$$

$$\text{(Log-moneyness width difference)} = \frac{\ln\left(K_{\max}(t,\tau)/S(t)\right) - |\ln\left(K_{\min}(t,\tau)/S(t)\right)|}{\sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \tag{22}$$

$$\text{(Strike price width log-ratio)} = \ln\left(\frac{K_{\max}(t,\tau) - S(t)}{S(t) - K_{\min}(t,\tau)}\right) = \ln\left(\frac{K_{\max}(t,\tau)/S(t) - 1}{1 - K_{\min}(t,\tau)/S(t)}\right), \tag{23}$$

$$\text{(Log-moneyness width log-ratio)} = \ln\left(\frac{\ln(K_{\max}(t,\tau)/S(t))}{|\ln(K_{\min}(t,\tau)/S(t))|}\right). \tag{24}$$

For all five definitions, if the asymmetry level is positive (negative), it means that the call (put)-side of integration domain is larger than the put (call)-side, while the zero level indicates that the integration domain is symmetric.

Figure 3 presents the histogram of the asymmetry level of integration domain for the maturity of two months. The figure shows that the asymmetry level is almost always negative regardless of how it is defined, and most of the positive levels are located near zero. This distribution suggests that put-side of integration domain is wider than the call-side in the S&P 500 index option market. Hence, it can be conjectured that if the integration domain symmetry leads to smaller truncation error of implied skewness estimator, there should be a negative relationship between the asymmetry level and the size of the truncation error, since an increase in asymmetry level does almost always mean a less negative value (which is closer to zero), rather than a more positive value (which is further away from zero).

Table 5 reports the regression result. The table suggests a positive relationship between the asymmetry level and the truncation error size, regardless of how the asymmetry level is defined. The positive relationship implies that truncation error tends to be smaller when the domain is more biased to OTM put side. This is consistent with the result in Section 4.2, given that the implied skewness is reported to be negative in the S&P 500 index options market. Another notable point is that the adjusted R^2 is larger when log-moneyness is used to define the asymmetry level. This again supports the idea that the log-moneyness of endpoint strike prices are important in explaining the relationship between truncation level and truncation error size.

6 A new approach to truncation error treatment

So far, this study has demonstrated how truncation affects implied moment estimators, and how the level of truncation relates to the size of truncation error. In addition, we show that the truncation error reduction effect of the existing truncation reduction methods, i.e., LE and DSym, is incomplete for the implied skewness and kurtosis estimators, and the magnitude of truncation error therefore can be significant even when those methods are employed. After investigating these issues, it is now natural to think about a new method which can control truncation error more effectively. We suggest that there is a need to tweak the question slightly to make it more realistic: Is there any way to reduce the "de facto" impact of truncation error on empirical analysis? In other words, there is a need to change the target from minimizing the truncation error itself to minimizing the impact of truncation error on empirical analysis.

If the model-free implied moment estimators are employed for making a cross-sectional comparison across the options on different underlying assets or tracking the time-series dynamics of the implied moments, there is a possibility of circumvention. As argued by Dennis and Mayhew (2002), if the size of estimation bias is roughly the same for all observations, one should be able to discern differences across the observations.[4] Hence, in this case, stabilizing the size of the truncation error can be an alternative objective of a truncation error treatment method. Based on this idea, we suggest a new truncation treatment method, i.e., DStab, which makes the size of the truncation error less volatile across the sample. Section 6.1 describes how to make the size of the truncation error less volatile by stabilizing the integration domain. Section 6.2 empirically tests whether the volatility of the truncation error in fact decreases by DStab.

6.1 Domain stabilization

In this subsection, we describe how the DStab works. This method suggests that with successful modeling of the relationship between the size of the truncation error (dependent variable) and some characteristics of integration domain (independent variables), the former becomes less volatile by stabilizing the latter. In Sections 4 and 5, the size of truncation error closely relates to the width and asymmetry level of the integration domain, and the relationship is strongest when measuring the

[4] Dennis and Mayhew (2002) adopt this argument as a rationale for not controlling strike price discreteness, arguing that the level of discreteness is roughly the same across all of their sample observations.

width and asymmetry level in a unit of volatility-adjusted log-moneyness. Given these findings, we suggest reducing the volatility of the truncation error size by stabilizing the width and asymmetry level of the integration domain in terms of volatility-adjusted log-moneyness. Specifically, DStab is conducted by setting the minimum and maximum volatility-adjusted log-moneyness values of daily integration domains as close to certain threshold values as possible.

To stabilize the minimum and maximum volatility-adjusted log-moneyness values, we need to trim some available OTM option prices as done by the DSym. Hence, DStab increases the size of truncation error in most cases, as shown in Section 4.2. The truncation error volatility does decrease, however, because this method heavily increases the size of relatively smaller truncation errors while leaving or only lightly increasing the relatively larger truncation errors. Given these characteristics, it can be argued that DStab results in a trade-off between the mean and volatility of truncation error size. Since there is a possibility that the implied moment estimate will convey misleading information if the truncation error is too large, we employ LE in conjunction with DStab to avoid exceedingly large truncation errors.

Two issues exist regarding the implementation of DStab. First, there is no definitive rule for setting the threshold level of volatility-adjusted log-moneyness at which to trim the OTM option prices. It is difficult to decide the threshold level because there is a trade-off between the better stabilization and the smaller truncation error, i.e., we need to discard more OTM option prices if we want to stabilize the truncation level more strongly via a more intensive trimming. To set a criterion, we first determine the percentage of daily integration domains that will be stabilized, and then choose the threshold values based on the percentage level. For instance, if we want the width and asymmetry level of 90 percent of the daily integration domains fixed, then we set the threshold levels at the 90^{th} percentile of the minimum volatility-adjusted log-moneyness and the 10^{th} percentile of the maximum volatility-adjusted log-moneyness for the corresponding maturity. Second, the level of implied volatility, what needs to be estimated, is required *ex ante*. Conducting the estimation in two stages circumvents this issue. Sections 4 and 5 show that the impact of truncation on the implied volatility estimator is much smaller than on the implied skewness or kurtosis estimator. Hence, we first estimate implied volatility without DStab but only with LE, and then take the implied volatility estimate for volatility adjustment.

Figure 4 illustrates the level of the volatility-adjusted endpoint log-moneyness of S&P 500 index options with the maturity of two months before and after DStab. DStab is conducted with two sets of threshold values to fix 90 and 99 percent of daily integration domains, respectively, in terms of volatility-adjusted endpoint log-moneyness. Figure 4b, as compared to Figure 4c, shows that the width of integration domain is larger in the case of the 90 percent stabilization than the 99 percent stabilization, but the volatility-adjusted endpoint log-moneyness level stabilizes more intensively in the case of the 99 percent stabilization. This again shows that one faces a trade-off when deciding the threshold values for DStab.

6.2 Effectiveness of domain stabilization

This subsection empirically tests whether DStab can reduce the volatility of truncation error size effectively, and compares the effectiveness of domain stabilization with some alternative methods. In addition to comparing the preliminary statistics before and after truncation treatment, we also conduct a set of variance comparison tests, using the test statistic of Levene (1960), which is robust to non-normality, and the two alternative statistics of Brown and Forsythe (1974) that are even more robust when dealing with skewed distributions. If a method effectively reduces the volatility of the truncation error size, the test statistics will indicate a significant decrease in variance after the employment of the method.

The alternative methods to which we compare DStab help achieve two goals. The first two alternative methods are employed to determine whether the two main features of DStab, i.e., the use of log-moneyness and volatility adjustment, enhance effectiveness. The first alternative method is another version of DStab for which the endpoint log-moneyness is replaced with the endpoint moneyness, while the second one is DStab without volatility adjustment. In addition, two versions of DSym are also employed as alternative methods to compare DStab to DSym. For each version of DSym, the integration domain symmetry is defined in terms of strike price and log-moneyness, respectively.

Table 8 reports the test results. Three interesting points can be found here. First, the size of truncation error becomes less volatile by DStab. In Panel A, the standard deviation of the proxy variable shows a decrease after DStab is employed. Furthermore, in Panel B, the variance comparison test result shows that the decrease in the standard deviation is statistically significant, both for the implied skewness and kurtosis estimators. Second, DStab becomes less effective if log-moneyness is not considered or volatility adjustment is not applied. Panel B shows that the decrease in the truncation error volatility is less significant if moneyness replaces log-moneyness. Furthermore, Panel B also reveals that the truncation error becomes more volatile after DStab for the implied kurtosis estimator if the volatility adjustment is not applied. Finally, both the mean and volatility of truncation error increase after DSym regardless of which moment is estimated or how the integration domain symmetry is defined. A possible reason for this is that while DSym fixes the asymmetry level of the integration domain, it does not consider the width. As mentioned in Section 4.2, the width of the integration domain closely relates to the size of the truncation error. Furthermore, truncation error is found to be smaller after DStab than after DSym. Panel A shows that the mean of truncation error size is smaller for either 90 or 99 percent stabilization when compared to DSym. This is because the number of discarded OTM option prices can set smaller for DStab in which the requirement of integration domain symmetry is relaxed.

Overall, the results in Table 8 show that DStab reduces the volatility of truncation error size effectively while minimizing the size of additional truncation error that is caused by further discarding OTM option price observations. Notably, this method shows significant effectiveness even when used for implied kurtosis estimation. Given that no truncation error treatment method has been suggested for implied kurtosis estimation thus far, reliability of the implied kurtosis estimator can increase by this method. Furthermore, DStab is also fould to be more effective for controlling the truncation error size of implied skewness estimate when compared to DSym, and this result makes DStab an even more attractive option for truncation error treatment.

7 Conclusion

Option prices are available only for a discrete and finite set of strike prices in most markets, and this limited availability is different from what is assumed in some models. Although it is relatively easy to mitigate strike price discreteness by using numerical techniques, it is more difficult to deal with the complete unavailability of extremely DOTM option prices, i.e., truncation. Given that truncation makes the implied moment estimators of Bakshi et al. (2003) biased, a proper treatment is needed to alleviate the estimation bias that is due to truncation. In particular, the implied skewness and kurtosis estimators, which rely more heavily on DOTM option prices, are more exposed to such errors, and therefore it can be conjectured that truncation should be treated more carefully for implied higher moment estimation.

This study investigates how effective DSym of Dennis and Mayhew (2002) is for mitigating the impact of truncation on the model-free implied moment estimators of Bakshi et al. (2003), and suggest how DSym can be improved so that it can mitigate the effect of truncation even when the implied RND is asymmetric. Our new method, DStab, is found to make the truncation error less volatile, whereas the error becomes more volatile after DSym.

With this result, we provide an empirical foundation for further studies on the higher moments of implied RND by answering questions on the impact of truncation on the implied skewness and kurtosis estimators of Bakshi et al. (2003). Especially, this study enables the higher moment estimators to be used with more reliability by proposing a new method of truncation treatment. It is especially encouraging that DStab is found to be effective for the implied kurtosis estimation, given that no method has thus far been suggested for controlling the impact of truncation on the simplied kurtosis estimator. In addition, given that the market tends to be less liquid for individual equity options, our method can provide a more reliable way in which implied kurtosis can be estimated for the individual equity options.

References

Bakshi, G., Cao, C., Chen, Z., 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003–2047.

Bakshi, G., Kapadia, N., Madan, D., 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bali, T. G., Murray, S., 2013, Does risk-neutral skewness predict the cross section of equity option portfolio returns?, *Journal of Financial and Quantitative Analysis* 48, 1145–1171.

Britten-Jones, M., Neuberger, A., 2000, Option prices, implied price processes, and stochastic volatility, *Journal of Finance* 55, 839–866.

Brown, M. B., Forsythe, A. B., 1974, Robust tests for the equality of variances, *Journal of the American Statistical Association* 69, 364–367.

Buss, A., Vilkov, G., 2012, Measuring equity risk with option-implied correlations, *Review of Financial Studies* 25, 3113–3140.

Carr, P., Madan, D., 2001, Optimal positioning in derivative securities, *Quantitative Finance* 1, 19–37.

Chang, B. Y., Christoffersen, P., Jacobs, K., 2013, Market skewness risk and the cross section of stock returns, *Journal of Financial Economics* 107, 46–68.

Chang, B. Y., Christoffersen, P., Jacobs, K., Vainberg, G., 2012, Option-implied measures of equity risk, *Review of Finance* 16, 385–428.

Chou, R. K., Chung, S. L., Hsiao, Y. J., Wang, Y. H., 2011, The impact of liquidity on option prices, *Journal of Futures Markets* 31, 1116–1141.

Christensen, B. J., Hansen, C. S., Prabhala, N. R., 2002, The telescoping overlap problem in options data, Working paper, University of Aarhus and University of Maryland.

Conrad, J., Dittmar, F., Ghysels, E., 2013, Ex ante skewness and expected stock returns, *Journal of Finance* 68, 85–124.

DeMiguel, V., Plyakha, Y., Uppal, R., Vilkov, G., 2013, Improving portfolio selection using option-implied volatility and skewness, *Journal of Financial and Quantitative Analysis* 48, 1813–1845.

Dennis, P., Mayhew, S., 2002, Risk-neutral skewness: Evidence from stock options, *Journal of Financial and Quantitative Analysis* 37, 471–493.

Diavatopoulos, D., Doran, J. S., Fodor, A., Peterson, D. R., 2012, The information content of implied skewness and kurtosis changes prior to earnings announcements for stock and option returns, *Journal of Banking and Finance* 36, 786–802.

Duan, J. C., Wei, J., 2009, Systematic risk and the price structure of individual equity options, *Review of Financial Studies* 22, 1981–2006.

Han, B., 2008, Investor sentiment and option prices, *Review of Financial Studies* 21, 387–414.

Jiang, G. J., Tian, Y. S., 2005, The model-free implied volatility and its information content, *Review of Financial Studies* 18, 1305–1342.

Kozhan, R., Neuberger, A., Schneider, P., 2013, The skew risk premium in the equity index market, *Review of Financial Studies* 26, 2174–2203.

Lee, G., 2015, Effectiveness of linear extrapolation in model-free implied moment estimation, Working paper, University of New South Wales.

Levene, H., 1960, Robust tests for equality of variances, in I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, and H. B. Mann, ed.: *Probability and Statistics: Essays in Honor of Harold Hotelling* (Stanford University Press: Menlo Park, CA).

Lin, B. H., Chang, I. J., Paxson, D. A., 2008, Smiling less at LIFFE, *Journal of Futures Markets* 28, 57–81.

Neumann, M., Skiadopoulos, G., 2013, Predictable dynamics in higher order risk-neutral moments: Evidence from the S&P 500 options, *Journal of Financial and Quantitative Analysis* 48, 947–977.

Xing, Y., Zhang, X., Zhao, R., 2010, What does the individual option volatility smirk tell us about future equity returns?, *Journal of Financial and Quantitative Analysis* 45, 641–662.

Table 1: Summary statistics of daily model calibration result

This table presents a set of summary statistics of the daily model calibration results and some related variables that are used for setting the model parameter values and generating simulated option prices. Panel A presents the Black-Scholes model calibration result. Panel B describes the stochastic volatility and jump model calibration result. Panel C compares the squared errors of the two model calibration results. Panel D provides information about the other variables that are used for generating simulated option prices.

	Mean	Standard error	5^{th} percentile	25^{th} percentile	Median	75^{th} percentile	95^{th} percentile
Panel A. Black-Scholes (BS) model parameter							
σ	0.2033	0.0622	0.1249	0.1530	0.2020	0.2315	0.3316
Panel B. Stochastic volatility and jump (SVJ) model parameters							
κ_v	4.4592	1.8378	2.0605	2.9532	4.2420	5.4808	8.0536
θ_v	0.2108	0.1248	0.0641	0.1220	0.1833	0.2662	0.4585
V_0	0.0494	0.0600	0.0088	0.0176	0.0342	0.0572	0.1422
σ_v	0.8116	0.3603	0.3564	0.5504	0.7700	0.9764	1.4665
μ_J	-0.1000	0.1301	-0.3193	-0.1964	-0.0886	-0.0006	0.0933
σ_J	0.1546	0.1136	0.0144	0.0555	0.1341	0.2358	0.3609
ρ	-0.6812	0.1114	-0.8718	-0.7566	-0.6760	-0.6044	-0.5060
λ	0.1583	0.1398	0.0150	0.0593	0.1207	0.2132	0.4562
Panel C. Squared error (SE)							
Sum of SE (BS)	16514.51	20298.67	3262.96	5209.51	8742.31	21507.79	49264.25
Sum of SE (SVJ)	876.24	3131.26	80.92	198.60	425.21	845.74	2576.11
SE per option (BS)	30.85	22.67	11.49	17.02	24.92	39.16	62.34
SE per option (SVJ)	1.98	8.29	0.26	0.54	0.96	1.90	6.11
Panel D. Other variables							
# of options	460.05	242.88	236	283	332	633	948
S&P 500 index	1183.21	189.91	860.02	1068.13	1179.21	1324.97	1491.56
3-month risk-free rate	0.0301	0.0205	0.0031	0.0119	0.0263	0.0505	0.0671
3-month dividend rate	0.0167	0.0051	0.0073	0.0137	0.0182	0.0205	0.0231

Table 2: Sample properties of S&P 500 index options dataset

This table presents a set of summary statistics of the S&P 500 index options data used in this study. The dataset spans an eleven-year time period from January 2000 through December 2010. Effective spread is defined as the difference between the ask quote and the mid-point. All statistics are collected after applying the following filtration conditions: (1) Observations with any missing data entry are removed; (2) observations are excluded if time to maturity is shorter than one week or longer than one year; (3) observations are included only if the sum of daily trading volume for the corresponding maturity date is nonzero—In other words, the entire daily observations for a maturity date are excluded if none of them are traded in that day; (4) observations are removed if the bid price is zero or higher than the ask price; (5) observations that violate the no-arbitrage restriction of upper and lower bounds are excluded; (6) observations with midpoints of bid and ask prices less than 0.375 are removed; and (7) observations with a bid-ask spread larger than the mid-point are excluded.

			Calls				Puts			
			# of trading days to expiration (t)			Sub-total	# of trading days to expiration (t)			Sub-total
Moneyness category	Moneyness ($m = K/S$)		$5 \leq t < 42$	$42 \leq t < 126$	$126 \leq t \leq 252$	# of obs.	$5 \leq t < 42$	$42 \leq t < 126$	$126 \leq t \leq 252$	# of obs.
	$m < 0.70$	Price	462.21	473.66	469.76		1.29	2.14	5.17	
		Eff. spread	1.37	1.35	1.42		0.42	0.42	0.57	
		BS implied vol.	0.9287	0.4874	0.3317		0.6358	0.4644	0.3847	
		# of obs.	7,658	7,027	8,153	22,838	7,592	24,873	27,541	60,006
ITM calls, OTM puts	$0.70 \leq m < 0.85$	Price	242.76	248.22	279.95		2.42	6.58	17.70	
		Eff. spread	1.35	1.36	1.39		0.40	0.57	0.92	
		BS implied vol.	0.4800	0.3199	0.2639		0.4097	0.3247	0.2707	
		# of obs.	26,402	34,399	30,033	90,834	35,449	47,189	31,580	114,218
	$0.85 \leq m < 1.00$	Price	88.05	110.26	145.87		9.78	25.63	47.75	
		Eff. spread	1.17	1.27	1.34		0.58	0.98	1.23	
		BS implied vol.	0.2496	0.2361	0.2227		0.2465	0.2390	0.2266	
		# of obs.	103,886	75,313	44,697	223,896	108,908	75,536	44,691	229,135
	$1.00 \leq m < 1.15$	Price	9.77	21.95	49.28		68.32	91.07	117.48	
		Eff. spread	0.59	0.90	1.21		1.24	1.33	1.37	
		BS implied vol.	0.1875	0.1861	0.1866		0.2117	0.2002	0.2012	
		# of obs.	80,673	73,663	42,739	197,075	73,099	62,471	35,690	171,260
OTM calls, ITM puts	$1.15 \leq m < 1.30$	Price	2.49	3.98	12.69		214.17	216.81	239.29	
		Eff. spread	0.57	0.55	0.82		1.62	1.70	1.72	
		BS implied vol.	0.3303	0.2074	0.1800		0.3866	0.2377	0.2106	
		# of obs.	6,791	21,343	25,948	54,082	15,275	15,379	13,443	44,097
	$m \geq 1.30$	Price	1.21	1.87	3.59		442.01	474.16	484.41	
		Eff. spread	0.59	0.55	0.57		1.71	1.79	1.94	
		BS implied vol.	0.4665	0.2843	0.2187		0.6716	0.3632	0.2705	
		# of obs.	892	5,182	15,984	22,058	10,669	13,640	12,704	37,013
	Total # of obs.					610,783				655,729
							1,266,512			

Table 3: Summary statistics of the truncation error proxy variable

This table presents a set of summary statistics of absolute percentage change in the implied moment estimate after linear extrapolation (APC), i.e.,

$$\mathrm{APC} = \left| \frac{(\text{Estimate after extrapolation}) - (\text{Estimate before extrapolation})}{(\text{Estimate before extrapolation})} \right|,$$

which is used as a proxy for the size of truncation error in Section 5. Since this variable can have an abnormally high value for the implied skewness estimate if the estimate originally has a near-zero value and its sign is switched after the linear extrapolation, we discard daily observations if the value of APC for the implied skewness estimate is larger than 1,000 percent. There are three such observations in our sample and, after this additional filtration, 2,750 daily observations remain.

Time period	Maturity	Moment	Mean	Median	Std. dev.	5^{th} pct.	95^{th} pct.	N
Entire sample period	2 months	Volatility	0.0146	0.0120	0.0104	0.0048	0.0337	2,750
		Skewness	0.0990	0.0770	0.1294	0.0112	0.2460	2,750
		Kurtosis	0.2131	0.1945	0.1140	0.0689	0.4197	2,750
	4 months	Volatility	0.0172	0.0120	0.0190	0.0042	0.0467	2,750
		Skewness	0.1338	0.0790	0.3244	0.0112	0.3371	2,750
		Kurtosis	0.2217	0.1855	0.1547	0.0576	0.5069	2,750
2000–2003	2 months	Volatility	0.0167	0.0137	0.0116	0.0051	0.0380	997
		Skewness	0.1386	0.0992	0.1917	0.0170	0.3344	997
		Kurtosis	0.2484	0.2254	0.1245	0.0810	0.4854	997
	4 months	Volatility	0.0193	0.0155	0.0142	0.0040	0.0474	997
		Skewness	0.1634	0.1238	0.1878	0.0150	0.3795	997
		Kurtosis	0.2691	0.2412	0.1450	0.0676	0.5394	997
2004–2007	2 months	Volatility	0.0147	0.0123	0.0100	0.0055	0.0327	999
		Skewness	0.0732	0.0676	0.0570	0.0069	0.1563	999
		Kurtosis	0.2040	0.1935	0.1043	0.0713	0.3846	999
	4 months	Volatility	0.0162	0.0109	0.0207	0.0048	0.0492	999
		Skewness	0.0780	0.0617	0.1436	0.0079	0.1540	999
		Kurtosis	0.1863	0.1652	0.1521	0.0571	0.3535	999
2008–2010	2 months	Volatility	0.0116	0.0093	0.0082	0.0039	0.0242	754
		Skewness	0.0806	0.0690	0.0705	0.0114	0.1699	754
		Kurtosis	0.1786	0.1649	0.0976	0.0591	0.3218	754
	4 months	Volatility	0.0158	0.0109	0.0217	0.0037	0.0451	754
		Skewness	0.1685	0.0736	0.5511	0.0172	0.3828	754
		Kurtosis	0.2062	0.1731	0.1551	0.0529	0.5008	754

Table 4: Relationship between the width of integration domain and the size of the truncation error

This table reports the regression results of the proxy variable for the truncation error size on the width of integration domain, which is defined in a number of different ways. A detailed description about the definition of asymmetry level can be found in Section 5.3. Volatility, which is the implied volatility estimate after linear extrapolation, is also considered as an independent variable in order to find out whether the improved explanatory power of volatility-adjusted log-moneyness width comes entirely from the level of implied volatility. In this table, ** and * denote statistical significance at the 1% and 5% levels, respectively.

Panel A. Implied skewness estimate

	2 months					4 months				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Strike price width	−0.0327 (−1.61)					−0.2810** (−6.78)				
Moneyness width		0.0312 (1.77)					0.0367 (0.95)			
Log-moneyness width			−0.0055 (−0.38)					−0.0849** (−2.87)		
Volatility-adjusted log-moneyness width				−0.0935** (−30.11)					−0.1834** (−25.64)	
Volatility					0.2505** (10.12)					0.7482** (10.96)
Constant	0.1139** (11.89)	0.0865** (11.58)	0.1014** (14.98)	0.5531** (36.31)	0.0422** (6.92)	0.3083** (11.65)	0.1140** (5.23)	0.1852** (9.77)	0.9816** (29.28)	−0.0407* (−2.39)
Adj. R^2	0.0006	0.0008	−0.0003	0.2478	0.0356	0.0161	0.0000	0.0026	0.1928	0.0416
N	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750

Panel B. Implied kurtosis estimate

	2 months					4 months				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Strike price width	−0.1526** (−8.67)					−0.2363** (−12.17)				
Moneyness width		−0.0543** (−3.50)					−0.0455* (−2.46)			
Log-moneyness width			−0.0697** (−5.55)					−0.1078** (−7.71)		
Volatility-adjusted log-moneyness width				−0.1080** (−45.29)					−0.1448** (−55.56)	
Volatility					0.1732** (7.90)					0.4224** (13.09)
Constant	0.2829** (33.98)	0.2348** (35.75)	0.2437** (41.14)	0.7379** (63.05)	0.1739** (32.12)	0.3685** (29.73)	0.2463** (23.70)	0.2871** (32.02)	0.8912** (72.93)	0.1232** (15.30)
Adj. R^2	0.0263	0.0041	0.0107	0.4272	0.0218	0.0508	0.0018	0.0208	0.5288	0.0584
N	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750

Table 5: Relationship between the asymmetry level of integration domain and the size of the truncation error

This table reports the regression results of the proxy variable for the truncation error size on the asymmetry level of integration domain, which is defined in a number of different ways. A detailed description about the definition of asymmetry level can be found in Section 5.4. Volatility-adjusted log-moneyness width, which is also shown in Table 4, is considered an independent variable used to control the impact of integration domain width on the truncation error size. In this table, ** and * denote statistical significance at the 1% and 5% levels, respectively.

	2 months					4 months				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Volatility-adjusted log-moneyness width	-0.0609^{**}	-0.0454^{**}	-0.0311^{**}	-0.0483^{**}	-0.0420^{**}	-0.1340^{**}	-0.1082^{**}	-0.0874^{**}	-0.0915^{**}	-0.0772^{**}
	(−13.79)	(−9.07)	(−5.31)	(−12.51)	(−10.70)	(−13.34)	(−9.98)	(−6.87)	(−10.46)	(−8.64)
Strike price width difference	0.0337^{**}					0.0589^{**}				
	(10.19)					(6.94)				
Moneyness width difference		0.0600^{**}					0.1055^{**}			
		(12.07)					(9.12)			
Log-moneyness width difference			0.0655^{**}					0.1059^{**}		
			(12.43)					(9.06)		
Strike price width log-ratio				0.1590^{**}					0.3195^{**}	
				(18.05)					(16.76)	
Log-moneyness width log-ratio					0.1819^{**}					0.3600^{**}
					(19.56)					(18.15)
Constant	0.4575^{**}	0.4121^{**}	0.3897^{**}	0.4471^{**}	0.4721^{**}	0.8381^{**}	0.7590^{**}	0.7424^{**}	0.7409^{**}	0.8049^{**}
	(25.92)	(21.81)	(19.67)	(28.74)	(31.76)	(21.41)	(18.48)	(17.55)	(21.16)	(24.28)
Adj. R^2	0.2749	0.2854	0.2876	0.3273	0.3395	0.2064	0.2162	0.2159	0.2674	0.2790
N	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750	2,750

Table 6: Effectiveness of the domain stabilization method

By reporting the sample mean and variance of the proxy variable for the size of the truncation error before and after treatment, as well as a set of variance comparison test statistics, this table shows how the mean and variance of the truncation error size are changed after applying a number of different truncation treatment methods. An n-percent domain stabilization is done by discarding the OTM option prices whose location on the corresponding integration domain is more left-sided than the n^{th} percentile of the left-side endpoint of the daily integration domains, or more right-sided than the $(100-n)^{\text{th}}$ percentile of the right-side endpoint of the daily integration domains for the corresponding maturity. On the other hand, domain symmetrization is done by discarding OTM option prices so that the width of integration domain is maximized while satisfying one of the following conditions:

$$\begin{cases} S - K_{\min} = K_{\max} - S, & \text{(Strike-price-based symmetrization)} \\ \ln(S/K_{\min}) = \ln(K_{\max}/S), & \text{(Log-moneyness-based symmetrization)} \end{cases}$$

where S is the underlying price, $K_{\min}$ is the minimum strike price after symmetrization, and $K_{\max}$ is the maximum strike price after symmetrization. In order to minimize the impact of outliers, a daily observation is discarded if the value of the proxy variable for the truncation error size, i.e., APC, is larger than 1,000 percent for any method or measure. The mean and standard deviation of the proxy variable are reported in Panel A. In Panel B, three different types of variance comparison test statistics are presented to show whether the change in the variance of the proxy variable is statistically significant. In Panel B, a (+) mark is placed together with the test statistic when variance is increased after the treatment, and a (−) mark is placed when variance is decreased. In this table, ** and * denote statistical significance at the 1% and 5% levels, respectively.

Panel A. Mean and standard deviation of truncation error							
Moment	Method	Measure	$\tau = 2$ months		$\tau = 4$ months		N
			Mean	Std. dev.	Mean	Std. dev.	
Skewness	No method applied		0.0988	0.1295	0.1272	0.2786	2,745
	90 percent stabilization	Volatility-adjusted log-moneyness	0.1351	0.0625	0.1932	0.0688	2,745
		Volatility-adjusted moneyness	0.1562	0.0677	0.2154	0.0766	2,745
		Log-moneyness	0.1019	0.0745	0.1205	0.0829	2,745
	99 percent stabilization	Volatility-adjusted log-moneyness	0.2633	0.0929	0.4544	0.2128	2,745
		Volatility-adjusted moneyness	0.2838	0.0900	0.4287	0.1793	2,745
		Log-moneyness	0.1191	0.0819	0.2968	0.2019	2,745
	Domain symmetrization	Strike price	0.5658	0.1931	0.5212	0.2803	2,745
		Log-moneyness	0.7925	0.2533	0.8383	0.4026	2,745
Kurtosis	No method applied		0.2130	0.1140	0.2204	0.1515	2,745
	90 percent stabilization	Volatility-adjusted log-moneyness	0.3473	0.0906	0.4274	0.0864	2,745
		Volatility-adjusted moneyness	0.3797	0.1015	0.4557	0.1076	2,745
		Log-moneyness	0.5094	0.3493	0.4825	0.3225	2,745
	99 percent stabilization	Volatility-adjusted log-moneyness	0.6408	0.0661	1.0458	0.0898	2,745
		Volatility-adjusted moneyness	0.6671	0.0733	1.0348	0.0820	2,745
		Log-moneyness	0.7033	0.4627	1.1414	0.6842	2,745
	Domain symmetrization	Strike price	0.6926	0.2232	0.6879	0.2873	2,745
		Log-moneyness	0.7944	0.2256	0.8478	0.2815	2,745

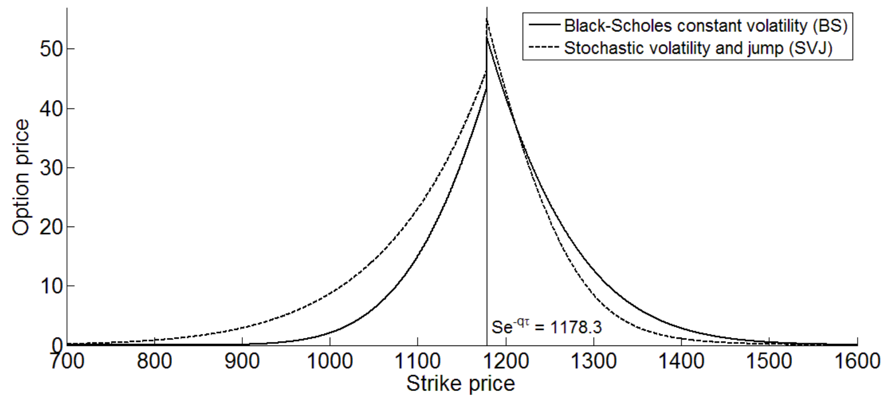
Table 8: Effectiveness of the domain stabilization method (*cont.*)

Panel B. Variance comparison test statistics (compared to when no method is applied)

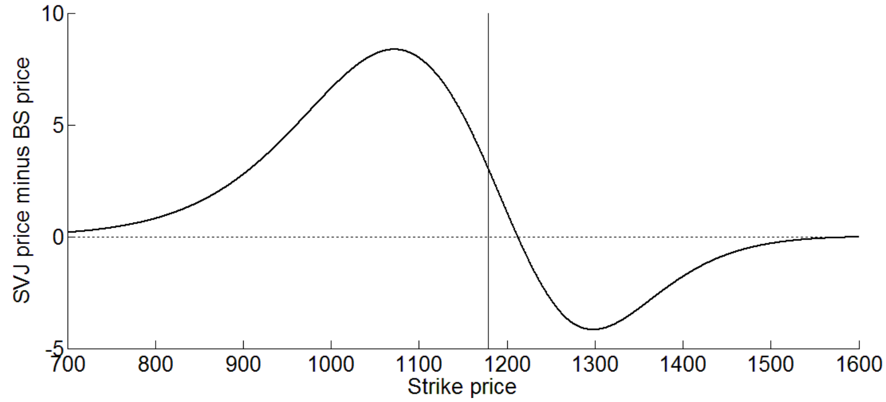
Moment	Method	Measure	Test statistics					
			$\tau = 2$ months			$\tau = 4$ months		
			Levene (1960)	Brown and Forsythe (1974)		Levene (1960)	Brown and Forsythe (1974)	
			Mean-centered	Median-centered	Trimmed-mean-centered	Mean-centered	Median-centered	Trimmed-mean-centered
Skewness	90 percent stabilization	Volatility-adjusted log-moneyness	(−) 8.90**	(−) 1.65	(−) 1.81	(−) 62.64**	(−) 24.89**	(−) 27.14**
		Volatility-adjusted moneyness	(−) 1.21	(−) 0.17	(−) 0.18	(−) 44.26**	(−) 15.28**	(−) 16.45**
		Log-moneyness	(−) 0.17	(−) 2.88	(−) 2.80	(−) 33.25**	(−) 8.91**	(−) 10.08**
	99 percent stabilization	Volatility-adjusted log-moneyness	(−) 56.78**	(−) 77.91**	(−) 77.89**	(−) 247.22**	(−) 317.29**	(−) 315.94**
		Volatility-adjusted moneyness	(−) 35.67**	(−) 52.19**	(−) 52.39**	(−) 100.09**	(−) 149.86**	(−) 148.30**
		Log-moneyness	(−) 5.09*	(−) 11.56**	(−) 11.56**	(−) 112.32**	(−) 129.28**	(−) 138.09**
	Domain symmetrization	Strike price	(+) 726.23**	(+) 736.61**	(+) 744.89**	(+) 156.12**	(+) 190.94**	(+) 191.83**
		Log-moneyness	(+) 1183.09**	(+) 1106.07**	(+) 1141.15**	(+) 577.59**	(+) 518.23**	(+) 564.96**
Kurtosis	90 percent stabilization	Volatility-adjusted log-moneyness	(−) 35.41**	(−) 25.63**	(−) 26.78**	(−) 175.06**	(−) 111.35**	(−) 122.32**
		Volatility-adjusted moneyness	(−) 1.73	(−) 0.54	(−) 0.57	(−) 46.55**	(−) 25.66**	(−) 27.96**
		Log-moneyness	(+) 776.63**	(+) 567.81**	(+) 609.75**	(+) 622.74**	(+) 494.56**	(+) 529.30**
	99 percent stabilization	Volatility-adjusted log-moneyness	(−) 269.97**	(−) 227.20**	(−) 236.04**	(−) 121.50**	(−) 75.78**	(−) 87.53**
		Volatility-adjusted moneyness	(−) 164.86**	(−) 135.12**	(−) 140.68**	(−) 219.86**	(−) 142.31**	(−) 156.24**
		Log-moneyness	(+) 1007.79**	(+) 695.38**	(+) 754.20**	(+) 1497.43**	(+) 1107.07**	(+) 1190.66**
	Domain symmetrization	Strike price	(+) 672.32**	(+) 652.02**	(+) 659.02**	(+) 530.97**	(+) 535.44**	(+) 539.51**
		Log-moneyness	(+) 645.48**	(+) 606.74**	(+) 621.49**	(+) 518.05**	(+) 504.26**	(+) 512.16**

Figure 1: Option price properties

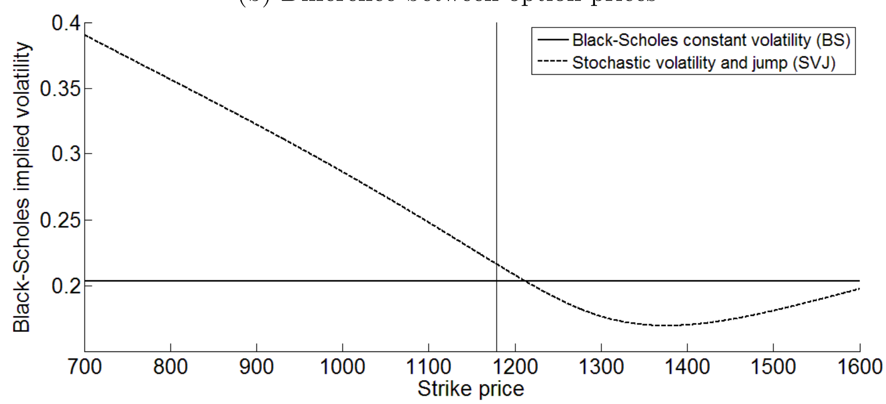
This figure illustrates a set of properties of the simulated option prices that are used in Section 4. Simulated option prices are generated using two option pricing models, i.e., Black-Scholes constant volatility (BS) and stochastic volatility and jump (SVJ) models, and the model parameter values that are determined based on the model calibration results on the S&P 500 index options dataset. The underlying price, risk-free rate, and dividend rate are set as the mean value for the same dataset. The time to maturity is set as three months to represent the maturities from two to four months, which are considered in Section 5. Figure 1a shows the option price level for strike price domain $[700, 1600]$. Figure 1b visualizes the option price difference between two models. Figure 1c demonstrates the Black-Scholes implied volatility curve for both models.



(a) Option prices



(b) Difference between option prices



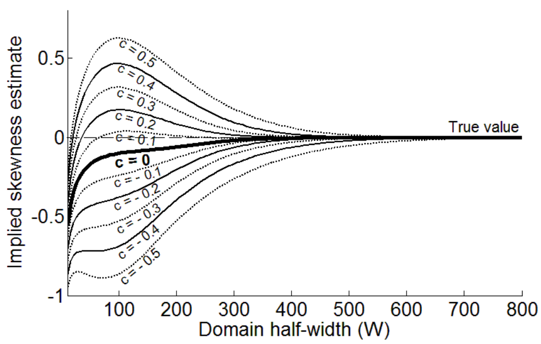
(c) Black-Scholes implied volatility curve

Figure 2: Asymmetry level of integration domain and the implied skewness estimate

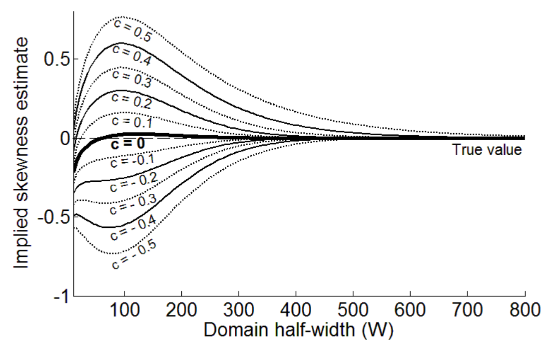
This figure illustrates the relationship between the asymmetry level of integration domain and the implied skewness estimate. In Figures 2a, 2c, and 2e, for each domain half-width W, the strike price domain of integration is set to be $[S-(1-c)W, S+(1+c)W]$, where $S = 1178.3$ is the dividend-free underlying price and c is the asymmetry coefficient. In figures 2b, 2d, and 2f, on the other hand, strike price domain of integration is set to satisfy the following conditions for each W:

$$(1+c)\ln(S/K_{\min}) = (1-c)\ln(K_{\max}/S); \quad \text{and} \quad K_{\max} - K_{\min} = 2W,$$

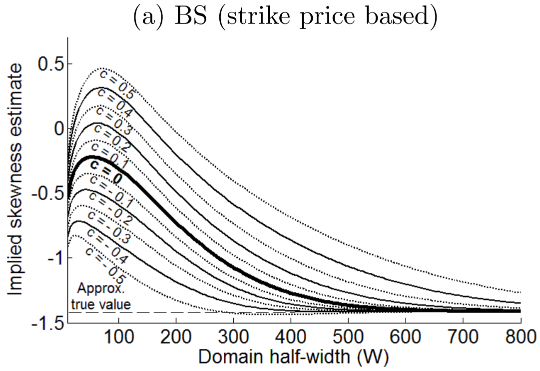
where $K_{\min}$ and $K_{\max}$ are the minimum and maximum strike prices, respectively. The Black-Scholes constant volatility (BS) model is used to generate option prices in Figures 2a and 2b. The stochastic volatility and jump (SVJ) model is used to generated option prices in Figures 2c–2f, while the sign of values for parameters μ and ρ are switched to be positive in Figures 2e and 2f to set the true implied skewness as positive. The straight line in each subfigure indicates the approximated true level of implied skewness, which is obtained by an estimation using the OTM option prices for the strike price domain $[3/S, 3S]$, where $S = 1178.3$ is the dividend-free underlying price. The strike price interval is fixed at 0.1.



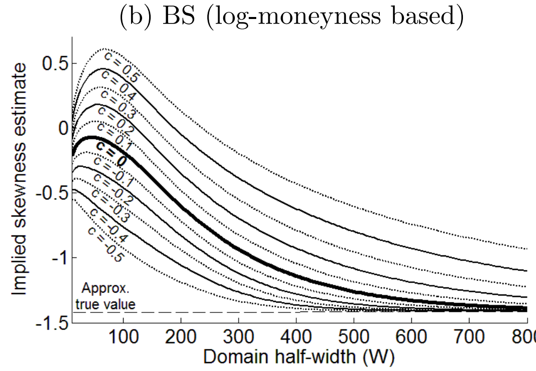
(a) BS (strike price based)



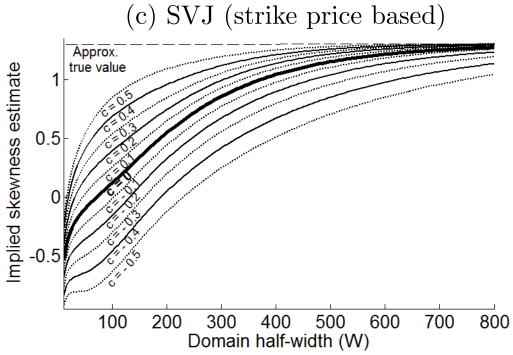
(b) BS (log-moneyness based)



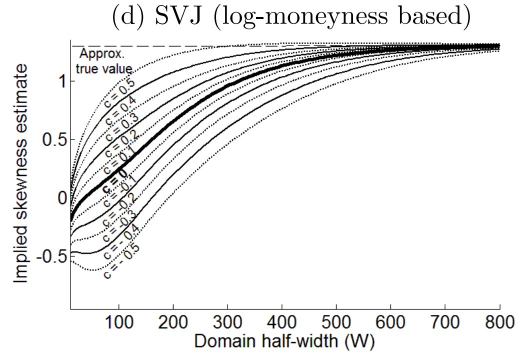
(c) SVJ (strike price based)



(d) SVJ (log-moneyness based)



(e) SVJ (sign switched, strike price based)



(f) SVJ (sign switched, log-moneyness based)

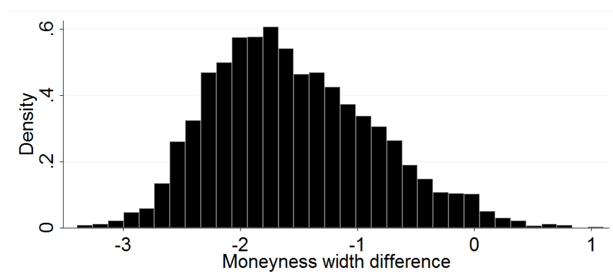
Figure 3: Sample distribution of domain asymmetry level

This figure presents the sample distribution of the asymmetry level of the integration domain level for the maturity of two months. The asymmetry level is defined in the following five different ways:

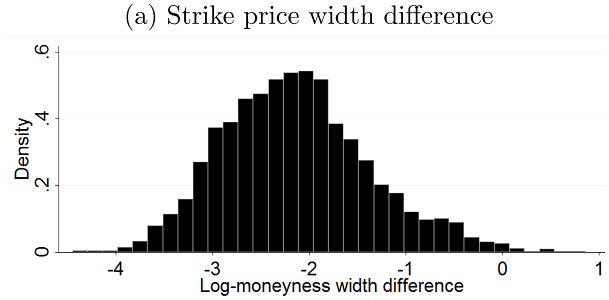
$$\begin{aligned}
(\text{Strike price width difference}) &= \frac{(K_{\max}(t,\tau) - S(t)) - (S(t) - K_{\min}(t,\tau))}{1000 \cdot \sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \\
(\text{Moneyness width difference}) &= \frac{(K_{\max}(t,\tau)/S(t) - 1) - (1 - K_{\min}(t,\tau)/S(t))}{\sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \\
(\text{Log-moneyness width difference}) &= \frac{\ln\left(K_{\max}(t,\tau)/S(t)\right) - |\ln\left(K_{\min}(t,\tau)/S(t)\right)|}{\sqrt{\tau} \cdot \mathrm{VOL}(t,\tau)}, \\
(\text{Strike price width log-ratio}) &= \ln\left(\frac{K_{\max}(t,\tau) - S(t)}{S(t) - K_{\min}(t,\tau)}\right) = \ln\left(\frac{K_{\max}(t,\tau)/S(t) - 1}{1 - K_{\min}(t,\tau)/S(t)}\right), \\
(\text{Log-moneyness width log-ratio}) &= \ln\left(\frac{\ln(K_{\max}(t,\tau)/S(t))}{|\ln(K_{\min}(t,\tau)/S(t))|}\right),
\end{aligned}$$

where $K_{\min}(t,\tau)$ and $K_{\max}(t,\tau)$ are the minimum and maximum strike prices for time t and maturity τ, respectively, $S(t)$ is the dividend-free S&P 500 index level at time t, and $\mathrm{VOL}(t,\tau)$ is the level of implied volatility for time t and maturity τ. Implied volatility level is estimated using Bakshi et al.'s (2003) implied volatility estimator and the linear extrapolation method. Linear extrapolation is applied up to the points where the strike prices are $3/S(t)$ and $3S(t)$, respectively.

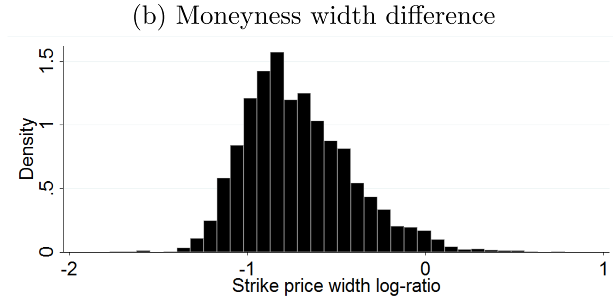
(a) Strike price width difference



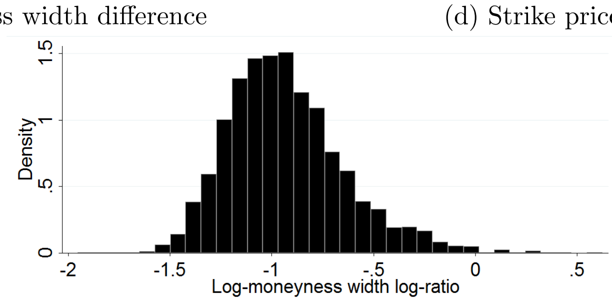
(b) Moneyness width difference



(c) Log-moneyness width difference



(d) Strike price width log-ratio



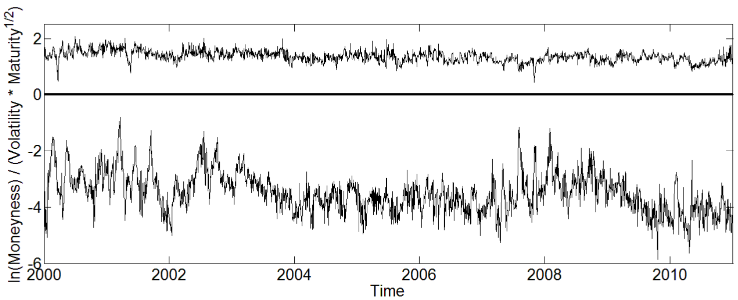
(e) Log-moneyness width log-ratio

Figure 4: Impact of domain stabilization

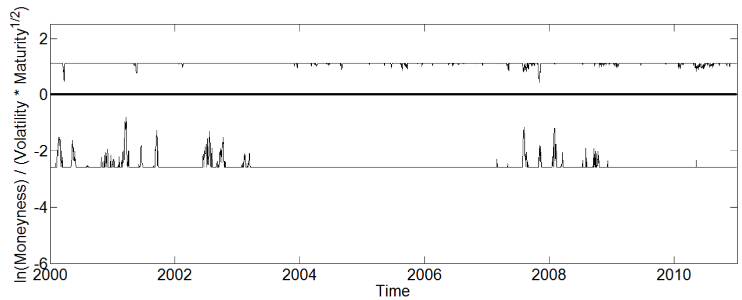
This figure illustrates the impact of domain stabilization by showing how the minimum and maximum volatility-adjusted log-moneyness values are changed after the stabilization. For the minimum and maximum strike prices at each time t, the volatility-adjusted log-moneyness is defined as

$$\frac{\ln(K/S(t))}{\mathrm{VOL}(t,\tau)\sqrt{\tau}},$$

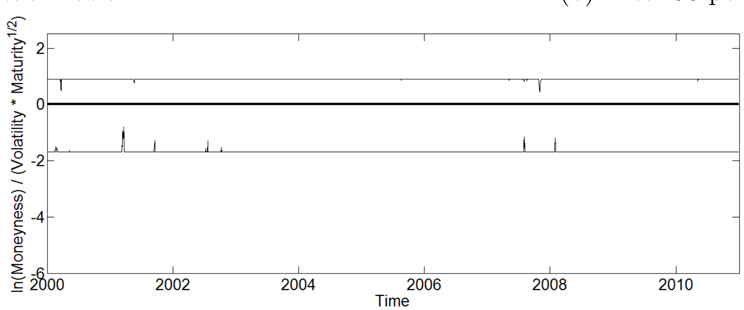
where K is the strike price, $S(t)$ is the dividend-free index level at time t, and $\mathrm{VOL}(t,\tau)$ is the implied volatility level for time t and maturity τ. The S&P 500 index options dataset, which spans the time period from January 2000 through December 2010, is used to generate daily implied volatility surfaces, from which the minimum and maximum volatility-adjusted log-moneyness values are extracted for the maturity of two months. An n-percent stabilization is done by discarding options whose volatility-adjusted log-moneyness is smaller than the n^{th} percentile of the minimum volatility-adjusted log-moneyness value of the daily integration domains, or larger than the $(100-n)^{\text{th}}$ percentile of the maximum volatility-adjusted log-moneyness value of the daily integration domains. The level of implied volatility is estimated using Bakshi et al.'s (2003) implied volatility estimator and the linear extrapolation method. Linear extrapolation is applied up to the points where the strike prices are $3/S(t)$ and $3S(t)$, respectively.



(a) Before stabilization



(b) After 90 percent stabilization



(c) After 99 percent stabilization