Using the Risk Neutral Density to Verify No Arbitrage in Implied Volatility

by Fabrice Douglas Rouah
www.FRouah.com
www.Volopta.com

Constructing implied volatility curves that are arbitrage-free is crucial for producing option prices that are sensible. In this Note we explain how the risk neutral density (RND) can be used to verify whether implied volatilities are arbitrage-free. The main references are Carr and Madan [2] and Fengler [3].

1 The Implied Volatility Curve and the RND

1.1 Constructing Implied Volatility and the RND

The steps to constructing an implied volatility curve and extracting the RND can be summarized as follows.

1. For a set of n strikes k_i, collect a set of market call prices c_i, all with the same maturity. Extract the implied volatility v_i from these call prices. This produces a set of triples $\{(k_i, v_i, c_i)\}_{i=1}^{n}$.

2. Expand the range of strikes and increase the granularity to produce a set of N strikes $\{K_i\}_{i=1}^{N}$, with increment dK. For example, if the market strikes span from \$20 to \$60 in increments of \$10, expand the range to span \$10 to \$80 in increments of \$0.05.

3. Select a curve-fitting method for the implied volatilities along the expanded strikes $\{K_i\}_{i=1}^{N}$. There are many choices for a volatility function, including but not limited to

   - Quadratic Deterministic Volatility Function (DVF).
   - Stochastic Volatility Inspired Model (SVI).
   - SABR Model.
   - Interpolation. If you use this choice you will need to specify values to extrapolate for the tails, such as flat extrapolation, for example.

   Estimate the parameters (if necessary) of the chosen volatility function, and fit the implied volatility V_i to each expanded strike. With the fitted volatility, obtain the fitted market price C_i using the Black-Scholes formula. This produces a set of N strike-price pairs $\{(K_i, C_i)\}_{i=1}^{N}$. The sequence of the data we need is thus

$$\{(k_i, c_i)\}_{i=1}^{n} \rightarrow \{(k_i, v_i)\}_{i=1}^{n} \rightarrow \{(K_i, V_i)\}_{i=1}^{N} \rightarrow \{(K_i, C_i)\}_{i=1}^{N}$$

4. Use the result of Breeden and Litzenberger [1] that the value at a price S^* of the discounted risk neutral density $f_{S_T}(S_T)$ is the second partial derivative of the call price with respect to strike, evaluated at S^*

$$e^{-rT} f_{S_T}(S^*) = \left. \frac{\partial^2 C}{\partial K^2} \right|_{K=S^*} \tag{1}$$

We use the set of pairs $\{(K_i, C_i)\}_{i=1}^N$ to obtain the RND at every strike K_i. This requires finite difference approximations. If we take central differences, then we will have $N-4$ points, since the first partial derivatives will be approximated by

$$dC_i = \left. \frac{\partial C}{\partial K} \right|_{K=K_i} \approx \frac{C_{i+1} - C_{i-1}}{2 \times dK} \quad \text{for } i = 2, \cdots, N-1 \tag{2}$$

while the second partial derivatives (the discounted RND) will be approximated by

$$e^{-rT} f_{S_T}(K_i) = \left. \frac{\partial^2 C}{\partial K^2} \right|_{K=K_i} \approx \frac{dC_{i+1} - dC_{i-1}}{2 \times dK} \quad \text{for } i = 3, \cdots, N-2 \tag{3}$$

1.2 Checking for Arbitrage

We need to verify that the RND produces a volatility curve that is arbitrage-free. Essentially, this means that the RND should be a true density, and that call prices obtained by numerical integration of the RND should show no arbitrage. There are two familes of tests:

1. Tests based on the RND.
2. Tests based on option strategies.

1.2.1 RND-Based Tests for Arbitrage

Tests based on the RND involve checking whether the RND is a true density (that is does not take on negative values and that it integrates to unity), and that call prices obtained by numerical integration of the RND free of arbitrage. Hence, using the RND to test for arbitrage means that

- We should be able to recover the original market call prices c_i by numerical integration of the RND.
- The RND should not take on negative values and it should integrate to unity.
- The RND should produce call prices that are decreasing monotonically in strike. That is, since the call price $C(K)$ with strike K is

$$C(K) = e^{-rT} \int_K^\infty (S_T - K) f_{S_T}(s)\, ds$$

we must have that the first derivative is negative for any strike K_1

$$\left.\frac{\partial C}{\partial K}\right|_{K=K_1} = -e^{-rT}\int_{K_1}^{\infty} f_{S_T}(s)\,ds < 0. \tag{4}$$

To verify this condition, at every market strike we take the finite difference approximation of $\frac{\partial C}{\partial K}$ from Equation (2) to obtain the left-hand side, we integrate the discounted RND to obtain the right-hand side, and we compare the two resulting quantities.

- The RND should produce call prices that are convex. That is, for any two strikes $K_1 < K_2$ the first derivative must increase in strike

$$\left.\frac{\partial C}{\partial K}\right|_{K=K_2} - \left.\frac{\partial C}{\partial K}\right|_{K=K_1} = e^{-rT}\int_{K_1}^{K_2} f_{S_T}(s)\,ds > 0. \tag{5}$$

To verify this condition we take our finite approximations in Equation (2) to obtain the left-hand side of Equation (5), we integrate the discounted RND to obtain the right-hand side, and we compare the two resulting quantities.

1.2.2 Tests Based on Option Strategies

These tests involve checking that option strategies that employ market prices of calls obtained with the fitted volatilities from Step 3 of Section 1.1, are sensible. For simplicity we assume that the expanded strikes $\{K_i\}_{i=1}^{N}$ are equally spaced. In other words, $dK_i = K_i - K_{i-1} \equiv dK$.

- Vertical call bull spreads. These spreads are long one call and short another call with a higher strike. A bull spread with strikes $K_{i-1} < K_i$ has price $C_{i-1} - C_i > 0$. The price Q_i of $\left(\frac{1}{dK}\right)$ units of vertical call bull spreads is therefore $Q_i = \frac{C_{i-1}-C_i}{dK}$. It is easy to show that at expiry[1], when the stock price is S_T, the bull spread has value in $[0,1]$, that is,

$$0 \leq \frac{(S_T - K_{i-1})^+ - (S_T - K_i)^+}{dK} \leq 1.$$

Taking expectations under the risk neutral measure and discounting, we obtain

$$0 \leq \frac{C_{i-1} - C_i}{dK} \leq e^{-rT}.$$

Hence, the value of the bull spread should lie in $\left[0, e^{-rT}\right]$ so that

$$Q_i = \frac{C_{i-1} - C_i}{dK} \in \left[0, e^{-rT}\right]. \tag{6}$$

[1] Indeed, when $S_T < K_{i-1}$ we have $Q_i = 0$, when $K_{i-1} < S_T < K_i$ we have $Q_i = \frac{S_T - K_{i-1}}{K_i - K_{i-1}} < 1$, and when $S_T > K_i$ we have $Q_i = \frac{K_i - K_{i-1}}{K_i - K_{i-1}} = 1$.

- Butterfly spreads. These spreads are long one call with strike strike K_{i-1}, short two calls with strike K_i, and long one call with strike K_{i+1}. A butterfly spread has price $C_{i-1} - 2C_i + C_{i+1}$. It can be shown that the price BS_i of $\left(\frac{1}{dK^2}\right)$ units of a butterfly spread approaches the Dirac delta function as $dK \to 0$. Its price for $dK > 0$ is

$$BS_i = \frac{C_{i-1} - 2C_i + C_{i+1}}{dK^2} \geq 0. \tag{7}$$

2 Illustration

We illustrate steps 1 through 7 above using a stock with spot price of \$423.19 and when the risk-free rate is 1%.

1. We collect strikes k_i and call prices c_i on $n = 22$ call options, with maturity 2 weeks. We extract the implied volatility v_i from each. The triples $\{(k_i, c_i, v_i)\}_{i=1}^n$ are listed in Table 1.

Table 1. Market Strikes (k_i), Call Prices (c_i), and Implied Volatility (v_i)

k_i	c_i	v_i	k_i	c_i	v_i
300	123.40	0.6842	410	21.08	0.4019
310	113.45	0.6566	420	14.92	0.3918
320	103.51	0.6197	430	9.93	0.3815
330	93.61	0.5928	440	6.42	0.3802
340	83.61	0.5311	450	3.62	0.3663
350	73.86	0.5205	460	2.06	0.3653
360	64.07	0.4838	470	1.22	0.3730
370	54.57	0.4669	480	0.62	0.3705
380	45.24	0.4420	490	0.39	0.3857
390	36.45	0.4267	500	0.17	0.3783
400	28.28	0.4120	510	0.13	0.4006

2. We expand the range of strikes from \$250 to \$600 in increments of $dK =$ \$0.025, which produces $N = 14,001$ strikes.

3. We select the DVF, SVI, and SABR models, along with linear, spline, and shape-preserving cubic interpolation. We obtain the following parameter estimates.

    - For the DVF we obtain

$$\beta_0 = 3.0376, \beta_1 = -0.01162, \beta_2 = 1.26475 \times 10^{-5}.$$

    - For the SVI model we obtain

$$a = -0.7096, b = 2.0331, \rho = -0.4654, m = -0.1699, \sigma = 0.4711.$$

- For SABR we obtain

$$\alpha = 7.8335, \rho = -0.2823, v = 2.6244.$$

The results are plotted in Figure 1. The models appear to provide a similar fit to the market volatilities. The extrapolation, however, is different, as expected.

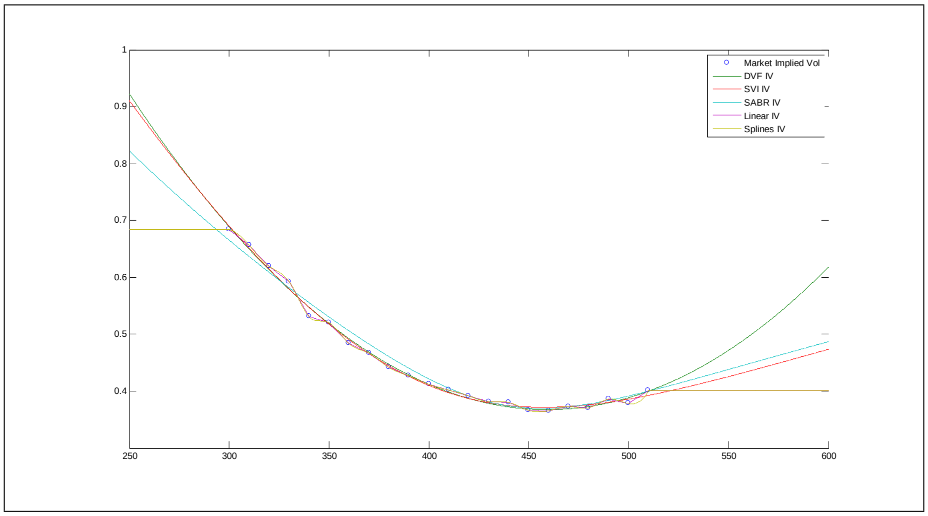

Figure 1. Market and Fitted Implied Volatilities

The call prices generated with these different volatility curves are plotted in Figure 2. At first glance, it seems that the choice of volatility curve is irrelevant since the call prices are all very close. We will see later that this is not the case, and that these call prices produce risk neutral densities that are vastly different.

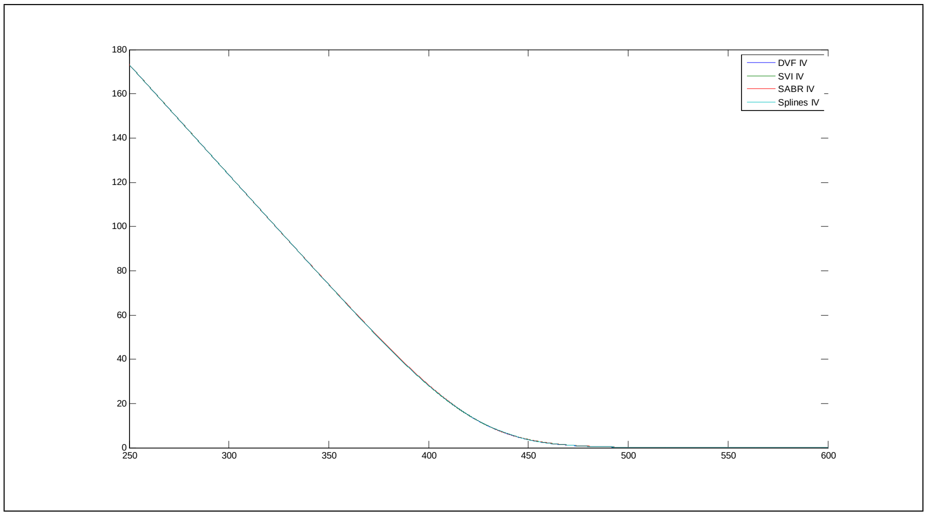

Figure 2. Call Prices From the Fitted Volatilities

4. We apply Equations (2) and (3) and obtain the RND f_{S_T} on $N - 4 = 13,997$ points. This appears in Figure 3. While the call prices in Figure 2 are very close, the risk neutral densities that are extracted from these prices are vastly different. The densities from the DVF, SVI, and SABR are reasonable, but the RND from splines is very jagged and takes on negative values (these have been floored at -0.01 to make the figure more presentable).

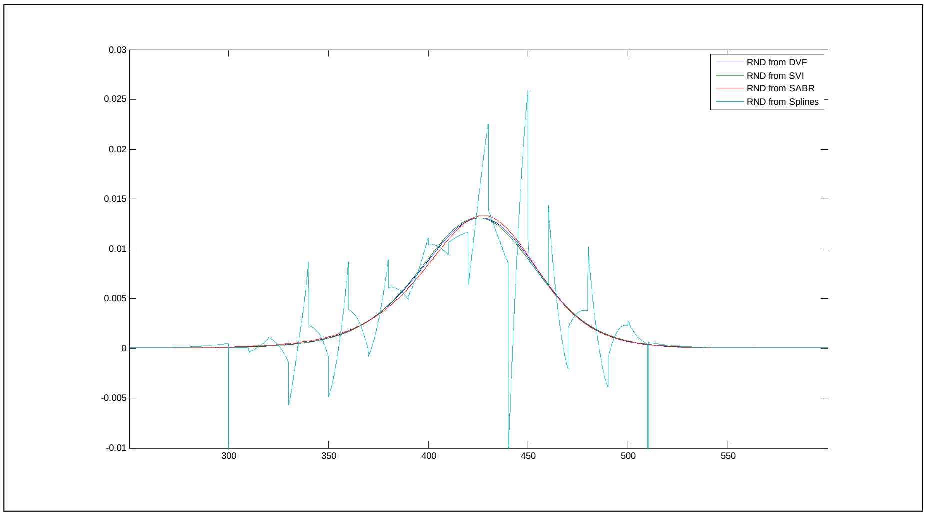

Figure 3. Risk Neutral Densities from the Fitted Call Prices

2.1 Checking for Arbitrage

We verify that the RNDs in Figure 3 are arbitrage free. We first calculate the area under the curve. This appears in the first row of Table 2.

Table 2. Arbitrage Statistics for the RNDs

Statistic	DVF	SVI	SABR	Linear	Splines	Cubic
Area under RND	1.0002	0.9993	0.9995	0.9999	0.9999	0.9999
Call pricing errors (%)	2.8561	-0.4368	0.8095	-0.0985	-0.1023	-0.1008
Eqn (4) errors (%)	-0.6758	0.0916	0.1196	2.2606	0.0503	0.0500
Eqn (5) errors (%)	0.0624	0.0611	0.0631	-88.1621	-3.0840	0.0736
% violation of En (6)	8.9	0	0	0	2.3	1.7
% violation of Eqn (7)	0	0	0	2.9	10.9	8.8

The table indicates that all methods produce RNDs that integrate close to unity. We next verify whether we can recover the original call prices from each of the RNDs. The prices appear in Table 3.

Table 3. Market Call Prices and Prices Recovered by RND

Strike	Market	DVF	SVI	SABR	Spline	Linear	Cubic
300	123.402	123.523	123.344	123.307	123.351	123.351	123.351
310	113.454	113.555	113.384	113.353	113.407	113.407	113.407
320	103.506	103.601	103.438	103.418	103.463	103.463	103.463
330	93.607	93.672	93.515	93.513	93.568	93.568	93.568
340	83.613	83.781	83.629	83.656	83.578	83.578	83.578
350	73.862	73.954	73.804	73.872	73.831	73.831	73.831
360	64.066	64.233	64.082	64.204	64.039	64.039	64.039
370	54.569	54.684	54.528	54.715	54.546	54.546	54.546
380	45.240	45.413	45.251	45.501	45.221	45.221	45.221
390	36.449	36.574	36.409	36.704	36.433	36.433	36.433
400	28.283	28.374	28.215	28.512	28.271	28.271	28.271
410	21.078	21.053	20.917	21.162	21.068	21.068	21.068
420	14.921	14.843	14.749	14.896	14.914	14.914	14.914
430	9.929	9.901	9.856	9.898	9.925	9.925	9.925
440	6.416	6.243	6.240	6.211	6.412	6.412	6.412
450	3.622	3.740	3.756	3.708	3.620	3.620	3.620
460	2.057	2.151	2.166	2.132	2.056	2.056	2.056
470	1.217	1.209	1.208	1.198	1.215	1.215	1.215
480	0.619	0.678	0.658	0.666	0.618	0.618	0.618
490	0.393	0.388	0.354	0.370	0.392	0.392	0.392
500	0.170	0.231	0.190	0.207	0.170	0.170	0.170
510	0.127	0.144	0.102	0.116	0.126	0.126	0.126

Clearly the models all recover the original market call prices adequately, especially for calls that are not too deep out-of-the money (recall that the spot price is \$423.19). This is confirmed by the average percent errors between the market call prices and the call prices produced by each of the implied volatility models. These errors appear in the second row of Table 2 and are all less than three percent in absolute value.

Next, we check the no-arbitrage condition that the first derivative of call prices with respect to strike is negative, in accordance with Equation (4). We evaluate the derivative at the market strikes, and present the results in Table 4 for the DVF, SVI, and SABR models, and in Table 5 for the splines, cubic, and linear interpolated models. The slopes and areas from the tables are generally comparable. The third row of Table 2 confirms that the average percentage error between the slope and area is less than one percent in absolute value for the DVF, SVI, SABR, and spline models, and just over two percent for the linear interpolated model. The results confirm that call prices are decreasing in strike. Note that since we are using central differences in Equation (2), the first and last strikes $K_1 = 300$ and $K_{22} = 510$ are excluded.

Table 4. Slope Checks for Arbitrage, Equation (4)

	DVF		SVI		SABR	
Strike	Slope	Area	Slope	Area	Slope	Area
310	-0.996	-0.996	-0.996	-0.995	-0.995	-0.995
320	-0.994	-0.994	-0.994	-0.994	-0.993	-0.992
330	-0.991	-0.991	-0.991	-0.991	-0.989	-0.989
340	-0.986	-0.986	-0.987	-0.986	-0.983	-0.983
350	-0.978	-0.978	-0.979	-0.978	-0.974	-0.973
360	-0.965	-0.965	-0.966	-0.965	-0.960	-0.959
370	-0.943	-0.943	-0.944	-0.944	-0.938	-0.937
380	-0.908	-0.909	-0.910	-0.909	-0.904	-0.903
390	-0.855	-0.856	-0.856	-0.856	-0.854	-0.853
400	-0.780	-0.780	-0.779	-0.779	-0.782	-0.781
410	-0.680	-0.680	-0.677	-0.677	-0.685	-0.685
420	-0.559	-0.559	-0.555	-0.554	-0.566	-0.565
430	-0.429	-0.429	-0.425	-0.424	-0.434	-0.434
440	-0.304	-0.305	-0.302	-0.302	-0.306	-0.306
450	-0.200	-0.200	-0.199	-0.199	-0.199	-0.199
460	-0.122	-0.122	-0.123	-0.123	-0.121	-0.121
470	-0.070	-0.070	-0.072	-0.072	-0.070	-0.070
480	-0.038	-0.039	-0.041	-0.041	-0.039	-0.039
490	-0.020	-0.021	-0.022	-0.022	-0.022	-0.022
500	-0.011	-0.012	-0.012	-0.012	-0.012	-0.012

Slope: left-hand side of Equation (4)
Area: right-hand side of Equation (4)

Table 5. Slope Checks for Arbitrage, Equation (4)

	Linear		Spline		Cubic	
Strike	Slope	Area	Slope	Area	Slope	Area
310	-0.994	-0.994	-0.996	-0.996	-0.995	-0.994
320	-0.994	-0.993	-0.991	-0.991	-0.993	-0.992
330	-0.988	-0.993	-0.996	-0.995	-0.991	-0.991
340	-1.000	-0.990	-0.991	-0.990	-0.984	-0.984
350	-0.968	-0.975	-0.972	-0.972	-0.971	-0.971
360	-0.976	-0.968	-0.972	-0.971	-0.966	-0.965
370	-0.937	-0.941	-0.938	-0.938	-0.941	-0.940
380	-0.918	-0.910	-0.914	-0.914	-0.909	-0.909
390	-0.852	-0.851	-0.849	-0.848	-0.851	-0.851
400	-0.779	-0.773	-0.775	-0.774	-0.772	-0.772
410	-0.671	-0.671	-0.667	-0.667	-0.671	-0.671
420	-0.559	-0.559	-0.565	-0.565	-0.559	-0.559
430	-0.439	-0.423	-0.417	-0.417	-0.412	-0.412
440	-0.297	-0.316	-0.316	-0.316	-0.300	-0.300
450	-0.226	-0.210	-0.220	-0.220	-0.198	-0.197
460	-0.122	-0.114	-0.106	-0.106	-0.120	-0.120
470	-0.065	-0.072	-0.075	-0.075	-0.075	-0.075
480	-0.044	-0.036	-0.035	-0.035	-0.042	-0.042
490	-0.017	-0.024	-0.023	-0.023	-0.026	-0.027
500	-0.015	-0.010	-0.015	-0.015	-0.013	-0.013

Slope: left-hand side of Equation (4)

Area: right-hand side of Equation (4)

We check the convexity condition in Equation (5). We present the results for the DVF, SVI, and SABR models for the market strikes in Table 6, and for the interpolated models in Table 7.

Table 6. Arbitrage Check for Convexity, Equation (5)

Strikes		DVF		SVI		SABR	
Lower	Upper	Diff	Area	Diff	Area	Diff	Area
310	320	0.0019	0.0019	0.0018	0.0018	0.0024	0.0024
320	330	0.0030	0.0030	0.0029	0.0029	0.0038	0.0037
330	340	0.0049	0.0049	0.0047	0.0047	0.0059	0.0059
340	350	0.0081	0.0081	0.0078	0.0078	0.0092	0.0092
350	360	0.0134	0.0134	0.0130	0.0130	0.0143	0.0143
360	370	0.0218	0.0218	0.0215	0.0215	0.0221	0.0221
370	380	0.0346	0.0346	0.0346	0.0346	0.0338	0.0338
380	390	0.0527	0.0527	0.0533	0.0533	0.0502	0.0502
390	400	0.0756	0.0756	0.0770	0.0770	0.0718	0.0718
400	410	0.1003	0.1003	0.1022	0.1021	0.0966	0.0966
410	420	0.1208	0.1208	0.1223	0.1222	0.1194	0.1193
420	430	0.1302	0.1302	0.1303	0.1302	0.1319	0.1319
430	440	0.1243	0.1243	0.1227	0.1227	0.1275	0.1275
440	450	0.1047	0.1046	0.1023	0.1022	0.1069	0.1068
450	460	0.0780	0.0779	0.0761	0.0760	0.0782	0.0782
460	470	0.0520	0.0519	0.0512	0.0511	0.0511	0.0511
470	480	0.0315	0.0315	0.0317	0.0317	0.0307	0.0307
480	490	0.0178	0.0177	0.0184	0.0184	0.0175	0.0175
490	500	0.0095	0.0095	0.0102	0.0102	0.0097	0.0097

Diff: left hand side of Equation (5)

Area: right hand side of Equation (5)

Table 6 indicates that the RND appear to satisfy the no-arbitrage requirement of Equation (5) for the DVF, SVI, and SABR models. In particular, the left- and right-hand sides of Equation (5) are almost always equal up to 4 decimal places, for every market strike. The last row of Table 2 indicates that the average percentage error between the left- and right- hands sides is less than 10 basis points. Note that since we are using central differences, only 19 of the 22 market strikes are represented. The results of Equation (5) for the interpolated models presented in Table 7.

Table 7. Arbitrage Check for Convexity, Equation (5)

Strikes		Linear		Splines		Cubics	
Lower	Upper	Diff	Area	Diff	Area	Diff	Area
310	320	0.000	0.002	0.005	0.005	0.002	0.002
320	330	0.006	-0.001	-0.004	-0.004	0.002	0.001
330	340	-0.012	0.003	0.005	0.005	0.007	0.007
340	350	0.032	0.015	0.019	0.018	0.013	0.013
350	360	-0.008	0.007	0.001	0.001	0.005	0.006
360	370	0.039	0.027	0.034	0.033	0.025	0.025
370	380	0.019	0.031	0.024	0.024	0.031	0.0317
380	390	0.066	0.059	0.066	0.065	0.058	0.058
390	400	0.073	0.078	0.074	0.074	0.079	0.079
400	410	0.108	0.102	0.108	0.108	0.101	0.101
410	420	0.112	0.112	0.102	0.102	0.112	0.112
420	430	0.120	0.135	0.148	0.149	0.146	0.147
430	440	0.142	0.108	0.102	0.101	0.113	0.113
440	450	0.070	0.105	0.095	0.096	0.102	0.103
450	460	0.104	0.096	0.114	0.114	0.077	0.077
460	470	0.057	0.043	0.031	0.031	0.046	0.045
470	480	0.021	0.035	0.040	0.041	0.033	0.034
480	490	0.026	0.012	0.012	0.011	0.015	0.015
490	500	0.002	0.014	0.008	0.009	0.014	0.014

Diff: left-hand side of Equation (5)

Area: right-hand side of Equation (5)

The entries in Table 7 and the percentage error in the last row of Table 2 both illustrate that the linear and spline interpolated volatilities in general do not satisfy the no-arbitrage condition. The areas and differences are not the same, and the values take on negative values. The cubic interpolation does a much better job.

Finally, for every set of call prices $\{C_i\}_{i=1}^N$ generated by each of the RNDs, we verify whether the vertical spreads $Q_i \in (0, 1)$ in accordance with Equation (6), and whether the butterfly spreads $BS_i \geq 0$, in accordance with Equation (7), are satisfied. The results of these tests appear in the last two rows of Table 2. They indicate only the SVI and SABR models produce call prices that pass the tests for all strikes.

To resume, while the linear and spline interpolated volatilities appear at first glance to provide an adequate model for a volatility curve, when the resulting call prices are subjected to rigorous arbitrage checks, they fail. The interpolated models all produce RNDs that take on negative values and the DVF model has trouble satisfying Equation (6). The SVI and SABR models are therefore the best choice of models for this set of data.

3 Illustration Using Interpolation Only

In this section we illustrate graphically that linear interpolation is a very poor choice of implied volatility function. Figure 5 presents fitted volatilities using linear interpolation, splines, and shape-preserving cubic splines, each with flat extrapolation. Both interpolation methods produce an implied volatility curve that is similar.

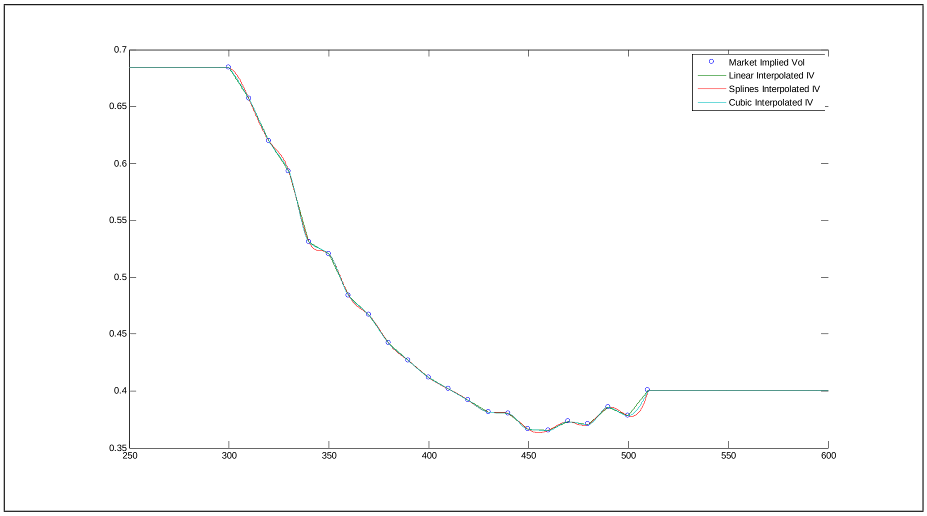

Figure 5. Linear and Splines Interpolation of Implied Volatilities

Figure 5a shows a close up of Figure 5, around a strike of 480. It is clear that the interpolation methods are not producing the exactly the same fitted volatilities.

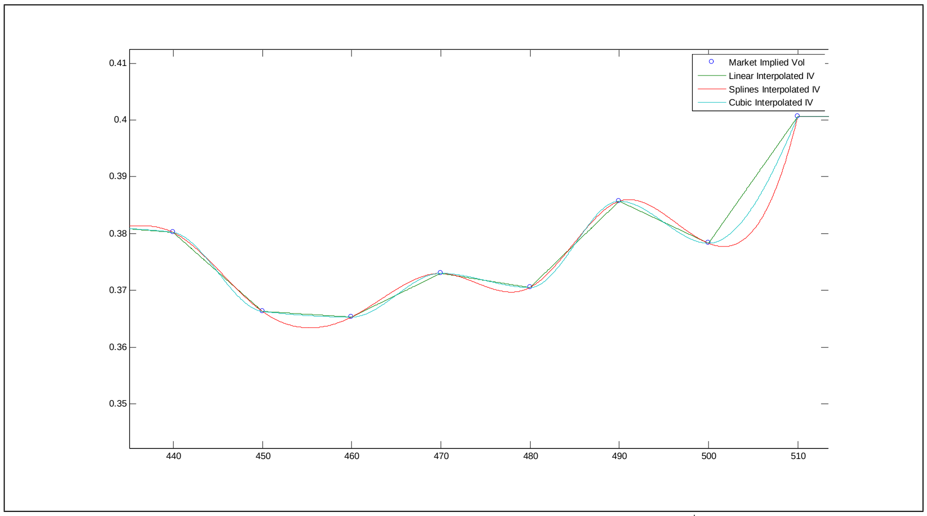

Figure 5a. Close up of Figure 5 Around $K = \$480$

The call prices from these two curves appear in Figure 6. Similar to Figure 2, the call prices are very close.

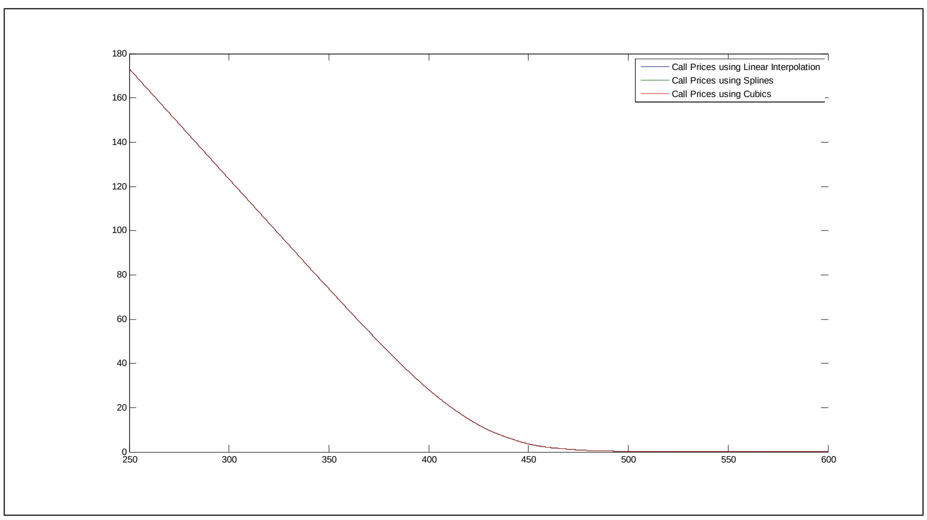

Figure 6. Call Prices from Interpolated Volatilities

Figure 6 is misleading, however, since the risk neutral densities extracted from these call prices are vastly different. This is illustrated in Figure 7, in which the RNDs have been floored at -0.01 and capped at 0.03 to make the figure more presentable.

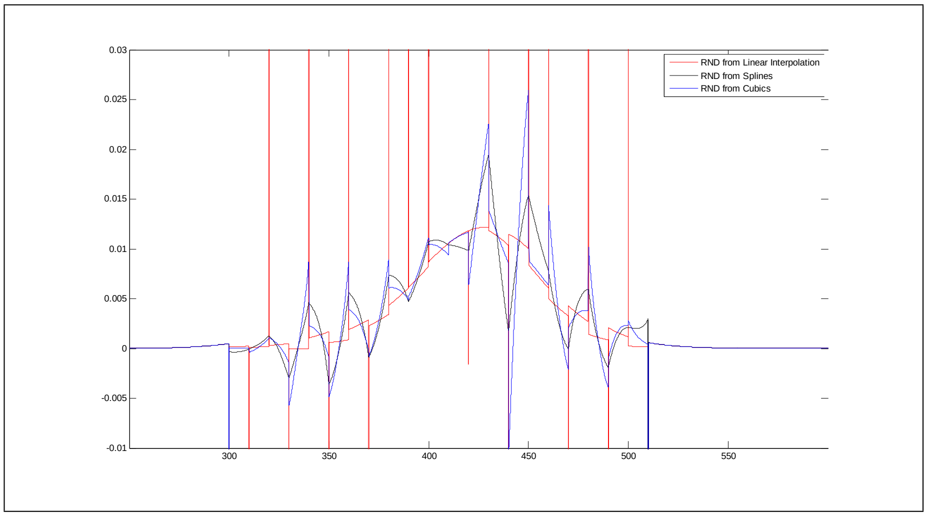

Figure 7. RNDs recovered from Interpolated Implied Volatilities

Figure 7 indicates that the RND from an implied volatility curve that is constructed by linear interpolation behaves erratically, and takes on negative values. The results of the previous section confirm that this is a poor choice of a volatility function because it leads to arbitrage.

References

[1] Breeden, D.T., and R.H. Litzenberger (1978). "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, Vol. 51, No. 4, pp. 621-651.

[2] Carr, P., and D.B. Madan (2005). "A Note on Sufficient Conditions for No Arbitrage," *Finance Research Letters*, Vol 2, pp. 125-130.

[3] Fengler, M.R. (2005). "Arbitrage-Free Smoothing of the Implied Volatility Surface," SFB 649 Discussion Paper, Sal. Oppenheim jr. & Cie, and Humboldt University.