Jesper Andreasen

Catch Up

We consider finite difference implementation of local volatility models when the underlying has a non-trivial drift. For this case, we develop a finite difference scheme that guarantees positive transition probabilities and also handle the case where the input option prices are not fully arbitrage consistent. We include C++ code for our finite difference solver.

National Media Museum from UK, No restrictions, via Wikimedia Commons

Bruno Dupire's Pricing with a Smile (1994) has had a lasting impact on the practical community. The local volatility model and its cousin the stochastic local volatility model are the industry standard for valuation of exotics. The attraction of the local volatility approach is that it allows direct calibration to vanilla option prices and the projection of their risks on vanilla options.

Another application of local volatility models is the reverse, where surfaces of local volatilities are used to parameterise arbitrage free surfaces of option prices. This methodology is useful for market making of exchange traded options.

Though the continuous time mathematics of the local volatility model is relatively straightforward, its actual numerical implementation is far from trivial. Actual implementation on computers requires some sort of discretization and it is far from trivial how this is done most effectively.

For local volatility models, it thus makes sense to choose a slower converging but more robust finite difference method

We are taught in skool to use second-order accurate finite difference methods, such as for example Crank-Nicolson, for the numerical solution of option pricing problems. That is generally a good approach because these methods are guaranteed to converge quickly. However, the convergence pattern of the Crank-Nicolson method exhibits small oscillations, and such oscillations are linked to "small" local arbitrages in the finite difference grid and wobble in the Greeks.

For local volatility models, convergence is not really an objective. Our main constraints are rather to hit the input option prices and avoid arbitrage. For local volatility models, it thus makes sense to choose a slower converging but more robust finite difference method.

Andreasen and Huge (2011a-b) analyze the fully implicit finite difference method. They show that the fully implicit finite difference method has non-negative transition probabilities everywhere and they derive a discrete version of the Dupire equation that enables exact calibration and full discrete consistency between the finite difference grid and the input option prices. But their analysis only applies to the case where the underlying has zero drift and/or cases where the drift can be eliminated by transformation.

Austing (2020) considers the local volatility model under non-trivial drift in a fully implicit finite difference scheme. However, his methodology somewhat intertwines drift and volatility, and it also doesn't guarantee non-negative transition probabilities everywhere in the finite difference grid.

Here, we describe a two-step fully implicit finite difference scheme that disentangles the drift and volatility calibration, as well as guarantees positive transition probabilities. As a further benefit, it facilitates so-called *catch-up-calibration* which remedies the case where input vanilla option prices are not fully consistent with absence of arbitrage.

The rest of the paper is organized as follows: the next section discusses cases where it isn't desirable to transform away the drift. We then discuss stability properties of different types of finite difference schemes before we introduce our two-step scheme and how it is calibrated, including the catch-up-calibration methodology. Finally, we give a summary of the full calibration algorithm.

Transformation

Consider a stock price s that evolves according to

$$ds = \mu\,(t,s)\,dt + \sigma\,(t,s)dW \tag{1}$$

where W is a Brownian motion under the risk neutral measure.

In many papers, including most of my own, the analysis is done under the assumption of zero drift. This is a convenient simplifying assumption and justifiable because often the drift can be transformed away. Consider for example the case where the drift is a linear function of spot, as for example

$$\mu = -q + \alpha s$$

where α is the difference between interest rate and dividend yield and q is the absolute dividend per time unit.

In this case, we can write

$$s(t) = e^{-\alpha t}s(0) - e^{-\alpha t}\int_0^t e^{\alpha u} q\,du + e^{-\alpha t}\tilde{s}(t) \qquad , d\tilde{s} = e^{\alpha t}\sigma dW \tag{2}$$

Hence, we can model the drift-less $\tilde{s}$ rather than the actual stock price.

There are, however, cases where such transformations are not desirable, and where it is advantageous to have explicit control over the placement of grid points in the spot space.

For example, for the pricing of continuously observed barrier options, it is convenient to place a grid point exactly on the barrier b and enforce absorption at this level, by setting $\mu(t,b) = \sigma(t,b) = 0$ for all t. This technique was used in the finite difference pricing of the barrier options in the previous issue, Andreasen (2022b). Not doing so will give a pricing inaccuracy and numerical noise of $O(\Delta s)$.

Another example is when we wish to align the finite difference grid explicitly to a range of strikes, for example, the strikes of exchange-traded equity option contracts. In this case, it is actually advantageous to place the grid points, not on the strikes, but symmetrically around them. This is also the case for barrier options that are observed at discrete time points, and digital options. Placing the grid points this way gives the highest accuracy. But more important, in a local volatility context, it increases stability of the solution, makes convergence graphs smoother, and improves the stability of the Greeks. To convince yourself that this is indeed the case, draw a little figure of numerical integration of a call payoff in two cases: Grid points on the strikes and around it. Placing the grid point on the strike gives an accuracy loss of order $O(\Delta s^2)$. But more importantly, not aligning the grid points to the strikes, on or around, and letting them float freely, will leave a noise term of $O(\Delta s^2)$ in the system. For digital payoffs, the loss of accuracy and numerical noise is $O(\Delta s)$.

Also worth mentioning here is that it is customary to use log transformation for the spot or pseudo spot, i.e., specify the finite difference grid in $\tilde{x} = \ln \tilde{s}$. This makes sense for many cases, but not necessarily here. The cost of using a logarithmic transform is that it is difficult to completely control the drift in the finite difference grid, i.e., ensure martingale conditions also hold in the finite difference grid. Instead, it seems to be a better approach to use a logarithmic *spacing* of the grid in s. This is the approach that we recommend for local volatility models.

There are, however, cases where such transformations are not desirable, and where it is advantageous to have explicit control over the placement of grid points in the spot space

Backwards and forwards

For the model (1), we have the continuous time and space backward equation

$$0 = \partial_t f + Af \qquad , A = \mu\partial_s + \frac{1}{2}\sigma^2\partial_{ss} \tag{3}$$

Here and in the following, we will generally assume zero discounting. Discounting is easy to include or apply externally for each time step. As such, it doesn't affect our analysis.

We wish to solve (3) on the discrete time line $0 = t_0 < t_1 < \dots$ and the spot grid $s_0 < s_1 < \dots < s_{n-1}$. The *theta* scheme for doing so is

$$\begin{aligned} f(t_{h+1/2}) &= [1 + (1-\theta)\Delta t\bar{A}]\,f(t_{h+1}) \\ [1-\theta\Delta t\bar{A}]\,f(t_h) &= f(t_{h+1/2}) \end{aligned} \tag{4}$$

where $\Delta t = t_{h+1} - t_h$, $\bar{A}$ is a discrete tridiagonal matrix approximation of the operator A, and θ is a constant in the interval $[0,1]$.

The system (4) is a backward system: we solve for the vector $f(t_h)$ given the vector $f(t_{h+1})$.

We will discuss $\bar{A}$ shortly but for now consider the time discretization, which is controlled by the parameter θ. θ is normally set to be one of the values:

$\theta = 0$: explicit scheme

Indeed, for local volatility models, the fully implicit scheme, $\theta = 1$, seems to be a better choice than Crank-Nicolson

$\theta = 1/2$: Crank-Nicolson scheme

$\theta = 1$: implicit scheme

θ determines convergence and stability properties in the time dimension.

The explicit scheme, aka the trinomial tree, $\theta = 0$, has an accuracy of $O(\Delta t)$ but it will only be stable, i.e., convergent, if $\Delta t = O(\Delta s^2)$. Ensuring that this stability condition is met requires many time steps, and on top of that it is a practical pain. Particularly so for local volatility. We don't consider this technique practical for the problem at hand.

The Crank-Nicolson scheme, $\theta = 1/2$, has an accuracy of $O(\Delta t^2)$ and it is von Neumann stable, in the sense that for all choices of $O(\Delta t, \Delta s)$, the solution will converge at the rate $O(\Delta t^2 + \Delta s^2)$, modulo the noise from grid alignment discussed in the previous section. If you're solving finite difference equations where the parameters of the model are given by physics, or even high authority, then the Crank-Nicolson method will yield superior accuracy. But even for $\mu = 0$, the transition probabilities inherent in the system (4) are not guaranteed by non-negative everywhere. This means that the numerical solution converges in an oscillatory manner. So if you're more tied to option prices than to convergence to continuous time equations, then Crank-Nicolson may not be the optimal numerical scheme.

Indeed, for local volatility models, the fully implicit scheme, $\theta = 1$, seems to be a better choice than Crank-Nicolson. The fully implicit scheme has an accuracy of $O(\Delta t)$. So a significant drop in formal accuracy relative to the Crank-Nicolson scheme. The attraction of the fully implicit scheme is that for $\mu = 0$, the transition probabilities are guaranteed to be non-negative everywhere. This means smooth, but slow, convergence, smooth Greeks, and it implies that vanilla option prices produced by the grid are always arbitrage-free everywhere. However, for this to hold when $\mu \neq 0$, additional trickery is needed.

Specifically, we define the first order *upwind* and *downwind* operators

$$\delta_s^+ f(s_i) = \frac{f(s_{i+1}) - f(s_i)}{s_{i+1} - s_i}$$
$$\delta_s^- f(s_i) = \frac{f(s_i) - f(s_{i-1})}{s_i - s_{i-1}} \qquad (5)$$

for interior points $0 < i < n-1$.

The second order operator can then be written as

$$\partial_{ss} f(s_i) = \frac{\delta_s^+ f(s_i) - \delta_s^- f(s_i)}{(s_i - s_{i-1})/2} \quad , 0 < i < n-1 \qquad (6)$$

The operator $\bar{A}$ is given by

$$\bar{A} f(s_i) = [\mu^+ \delta_s^+ - \mu^- \delta_s^- + \frac{1}{2}\sigma^2 \partial_{ss}]\, f(s_i) \quad , 0 < i < n-1 \qquad (7)$$

on the interior points. Here $\mu^+ = \max(0, \mu)$ and $\mu^- = -\min(\mu, 0)$.

The system (4) becomes a tridiagonal matrix system that is closed by setting

$$\bar{A} f(s_i) = 0 \quad , i = 0, n-1 \qquad (8)$$

on the boundaries. That is, we have absorption at the boundaries.

The point here is that the first order difference goes in the direction of the drift μ. When μ is positive, we take the first order derivative upwards, and when the drift is negative, we go downwards. This, in combination with the fully implicit scheme, can be proven to guarantee positive transition probabilities everywhere in the finite difference grid.

Using up and down-winding, rather than the more conventional central differencing for the first order operator, means that the accuracy of the scheme drops from $O(\Delta s^2)$ to $O(\Delta s)$. This has little practical consequence as, ultimately, the scheme will be calibrated exactly to option prices and forwards.

The positive transition probability property can be proven using next generation matrix algebra as in Andreasen (2022a), but it is more insightful to actually look at the numbers. To do so, we note that density for the backward system (4) solves the *forward* system

$$[1 - \theta \Delta t \bar{A}]' \, p(t_{h+1/2}) = p(t_h) \qquad (9)$$
$$p(t_{h+1}) = [1 + (1-\theta)\Delta t \bar{A}]' \, p(t_{h+1/2})$$

where ' denotes transpose. A derivation is given in Andreasen (2022a).

Equations (9) constitute a forward system where we solve for $p(t_{h+1})$ given $p(t_h)$. The initial boundary condition is $p(t_0, s_i) = 1_{s_i = s(t_0)}$. Note how time and order of solution is flip-flopped relative to the backward system (4). As such, (9) is the discrete Fokker-Planck equation for the system (4).

Figures 1, 2, and 3 show transition probabilities for different number of periods for the explicit, the Crank-Nicolson, and the fully implicit scheme. All calculations are done without drift and under the same grid dimensioning. Table 1 shows the resulting option prices when integrating the transition probabilities over a call option payoff.

Key observations here include:

- The explicit scheme goes increasingly bananas as we move forward in time. If we use sufficiently small time steps, then this can be prevented.
- The fully implicit scheme produces positive transition probabilities. For constant volatility, the one-step solution is a Laplace density.
- The Crank-Nicolson scheme produces transition probabilities that are non-negative, but they die out as we take more time steps.

- The Crank-Nicolson method can be seen as a mix of the explicit and the fully implicit methods: over the first half time step we go explicit, over the second we go implicit. The one period transition probability is thus a mix of the transition probabilities of the explicit and implicit method. We can eliminate the negative transition probability phenomenon in the Crank-Nicolson scheme if we take small time steps — as in the explicit method.

- Surprisingly, and despite the oscillatory nature of the transition probabilities under Crank-Nicolson discretization, the resulting finite difference prices are significantly more accurate, and in fact are very close to the closed form price with only a few time steps in the numerical solution.

Figure 4 shows transition probabilities for the fully implicit method when $\mu < 0$ and low volatility, for the case of central differencing, respectively, up and down-winding. Winding for the first order derivative eliminates the negative transition probabilities.

When transition probabilities are non-negative everywhere, the finite difference scheme has a probabilistic interpretation as a discrete time and state Markov chain.

Non-negative transition probabilities every in the grid guarantee arbitrage free option prices for all expiries and strikes. Figure 5 shows a snap shot of a finite difference generated implied volatility smile relative to call and put bids and offers on the exchange.

The repository, Andreasen, Huge and Kryger-Baggesen (2022), contains C++ code for the finite difference solver that we have used. We encourage readers to check it out and use it for making their own experiments.

When transition probabilities are non-negative everywhere, the finite difference scheme has a probabilistic interpretation as a discrete time and state Markov chain

2-Step

We have seen that combining the fully implicit method with up and down-winding for the drift produces positive transition probabilities everywhere. However, drift and volatility are still somewhat entangled, and calibration of drift and volatility will potentially need to be simultaneous.

To avoid this, consider the following backward scheme

$$
\begin{aligned}
&[1-\Delta t\tfrac{1}{2}\sigma^2\partial_{ss}]\, f(t_{h+1/2}) = f(t_{h+1}) \\
&[1-\Delta t\,(\mu^+\delta_s^+ - \mu^-\delta_s^-)]\, f(t_h) = f(t_{h+1/2})
\end{aligned}
\tag{10}
$$

The scheme (10) is in reality a split scheme where we do the drift and volatility operators in two separate time steps. This technique is often applied for multi-dimensional solvers where each dimension is handled in separate time steps. But along the lines of what we discussed in the previous section, you can also view the Crank-Nicolson scheme as a split scheme: first explicit, then implicit.

We implement (10) with absorption at the boundaries, i.e.,

$$\mu(s_i) = \sigma(s_i) = 0 \quad , i = 0, n-1 \tag{11}$$

The associated forward equation for the density is

$$
\begin{aligned}
&[1-\Delta t(\mu^+\delta_s^+ - \mu^-\delta_s^-)]'\, p(t_{h+1/2}) = p(t_h) \\
&[1-\Delta t\tfrac{1}{2}\sigma^2\partial_{ss}]'\, p(t_{h+1}) = p(t_{h+1/2})
\end{aligned}
\tag{12}
$$

With initial European option prices given by

$$c(t_h, s_j) = \sum_{i=0}^{n-1} p(t_h, s_i)\,(s_i - s_j)^+ = \sum_{i=j}^{n-1} p(t_h, s_i)\, s_i - s_j \sum_{i=j}^{n-1} p(t_h, s_i) \tag{13}$$

we obtain the discrete Dupire equation

$$[1-\Delta t\tfrac{1}{2}\sigma^2\partial_{ss}]\, c(t_{h+1}) = c\,(t_{h+1/2}) \tag{14}$$

See Andreasen (2022a) for a derivation.

We note that the drift will only affect the expectation of the stock over the first half time step. The stock will always be a martingale over the second half time step. On the other hand, the volatility specification doesn't affect the expectation of the stock. This splits the calibration in two steps:

(10b) is used to calibrate the drift, i.e., to match the forward over the time step $[t_h, t_{h+1/2}]$.

(14) is used over the time step $[t_{h+1/2}, t_{h+1}]$ to calibrate the local volatilities to European option prices.

For the latter, we note that using up and down-winding adds a bit of diffusion to the solution. This has to be incorporated into the calibration. We give the details in the next sections.

Hitting the forwards

Let

$$g(t_h, s_i) = E[s\,(t_{h+1}) \mid s\,(t_h) = s_i] \tag{15}$$

be the one period forward price.

We note that due to absorption, we have at the boundaries

$$g(t_h, s_0) = s_0$$

$$g(t_h, s_{n-1}) = s_{n-1} \tag{16}$$

For interior points, we can choose g freely and, for example, give it the linear specification in (2):

$$g(t_h, s_i) = e^{-\alpha\Delta t} s_i - q\Delta t \qquad ,0 < i < n-1 \tag{17}$$

A more sophisticated specification can also be used.

If $\{G(T)\}_{T\geq 0}$ are the initial forward prices, then we need g to satisfy

$$p(t_h, s_0)\, s_0 + \sum_{i=1}^{n-2} p(t_h, s_i)\, g(t_h, s_i) + p(t_h, s_{n-1})\, s_{n-1} = G\,(t_{h+1}) \tag{18}$$

We can adjust one parameter of the one period forward specification $g(t_h, s_i)$ so (16) is met. Once we have the function $g(t_h, s_i)$ we can insert it into (9b) to get

$$[1 - \Delta t(\mu^+ \delta_s^+ - \mu^- \delta_s^-)]\, g(t_h, s_i) = s_i \tag{19}$$

This can be solved for $\mu(t_h, s_i)$ on the interior points:

$$\mu(t_h, s_i) = \frac{(g(t_h, s_i) - s_i)^+}{\Delta t \delta_s^+ g(t_h, s_i)} - \frac{(s_i - g\,(t_h, s_i))^+}{\Delta t \delta_s^- g(t_h, s_i)} \qquad ,0 < i < n-1 \tag{20}$$

This is a well-defined solution as long as g is strictly increasing in s, hence we require that the lower and upper bounds of the grid are chosen so that

$$s_0 < g(t_h, s_i) < \ldots < g(t_h, s_{n-2}) < s_{n-1} \tag{21}$$

This is not a very strict restriction.

Good to catch up

We can now use equation (14) to fit the model to the option prices. Specifically, if C are the *input* option prices and c are the model option prices, then we set

$$\sigma(t_h, s_i)^2 = \min(\max(\underline{\sigma}^2, 2\frac{C(t_{h+1}, s_i) - c(t_{h+1/2}, s_i)}{\Delta t \delta_{ss}\, C(t_{h+1}, s_i)}), \bar{\sigma}^2) \tag{22}$$

Using $c(t_{h+1/2}, s_i)$ in the Dupire equation (22) compensates for the fact that the up and down-winding will introduce a bit of additional volatility in the finite difference solution.

Bounding of the local variance in (22) means that we can handle the case when input option prices are not fully arbitrage consistent. It is important to note that the chop-chop only has a local effect. As we're using the half step option price $c(t_{h+1/2}, s_i)$ in (22), the calibration procedure will *catch-up* after local problems with arbitrage in subsequent time steps.

I call this methodology *catch-up-calibration*. I have a hunch that I'm not the only one who has known about this trick for a while. The kool thing about it is that it has applications for other problems, such as for example model calibration by Monte-Carlo. But that's a story for another day.

So is the fact that the Babylonian algorithm, discussed two issues ago in Andreasen (2022a), can be extended to the case of drift using the scheme and algorithms presented in this paper.

Algorithm summary

Here's a summary of the calibration algorithm:

0. Set $h = 0$.
1. Use equation (18) to calibrate one parameter of the one period forward function $\{g(t_h, s_i)\}_{i=1,\ldots,n-2}$ so that we hit the initial forward $G(t_{h+1})$.
2. Compute the corresponding drift $\{\mu(t_h, s_i)\}_{i=1,\ldots,n-2}$ using (20).
3. Roll forward using (12a) to produce $\{p(t_{h+1/2}, s_i)\}_{i=0,\ldots,n-1}$.
4. Calculate the half time step option prices $\{c(t_{h+1/2}, s_i)\}_{i=0,\ldots,n-1}$ using (13).
5. Compute local volatility using the catch-up equation (22).
6. Roll forward using (12b) to produce $\{p(t_{h+1}, s_i)\}_{i=0,\ldots,n-1}$.
7. Set $h := h+1$ and go to 1.

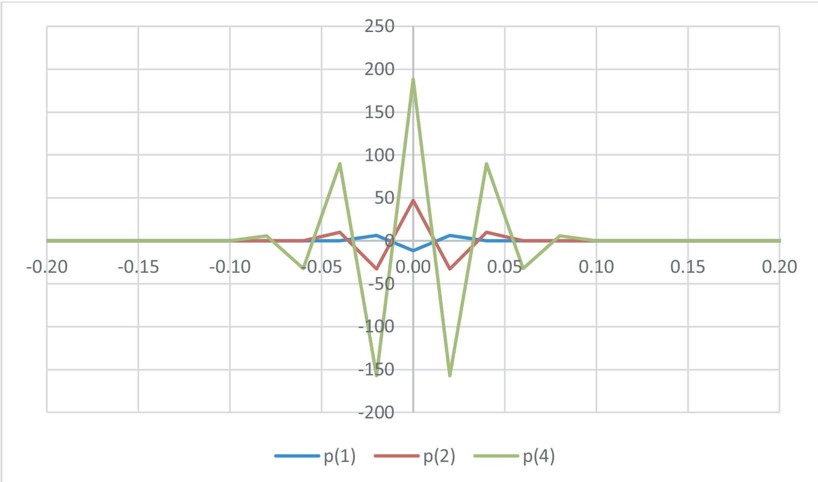
About the Author

Jesper Andreasen is the Kwant Daddy at Saxo Bank in Copenhagen.
Email: kwant.daddy@saxobank.com

REFERENCES

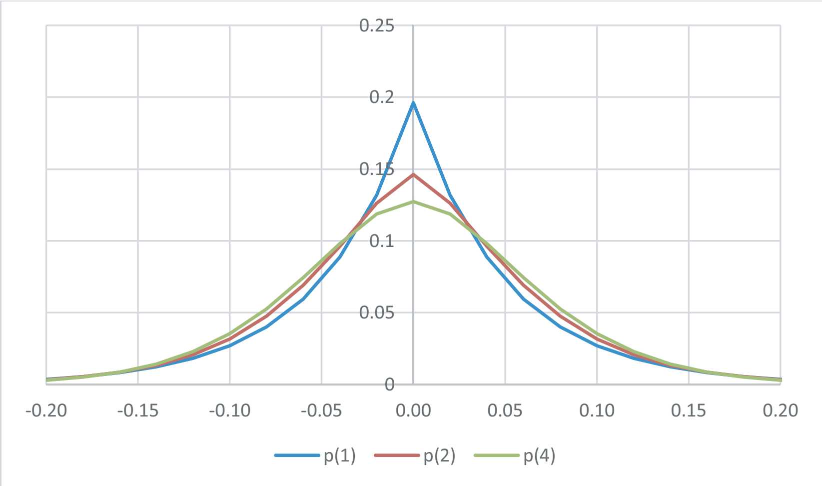
Andreasen, J. 2022a. By the Rivers of Babylon. *Wilmott* September.
Andreasen, J. 2022b. Three Strikes and You're Out. *Wilmott* November.
Andreasen, J. and Huge, B. 2011a. Volatility Interpolation. *Risk* March.
Andreasen, J. and Huge, B. 2011b. Random Grids. *Risk* June.
Andreasen, J., Huge, B. and Kryger-Baggesen, F. 2022. https://github.com/brnohu/catch_up
Austing, P. 2020. Finite Difference Schemes with Exact Recovery of Vanilla Option Prices. *Risk* November.
Dupire, B. 1994. Pricing with a Smile. *Risk* January.

Figure 1: Transition Probabilities for Explicit Scheme.



Transition probabilities from $s(0) = 0$ to $s(0.5)$ with 1, 2, and 4 time steps in explicit scheme $\theta = 0$. Parameters are $\mu = 0$, $\sigma = 0.1$.

Figure 2: Transition Probabilities for Implicit Scheme.



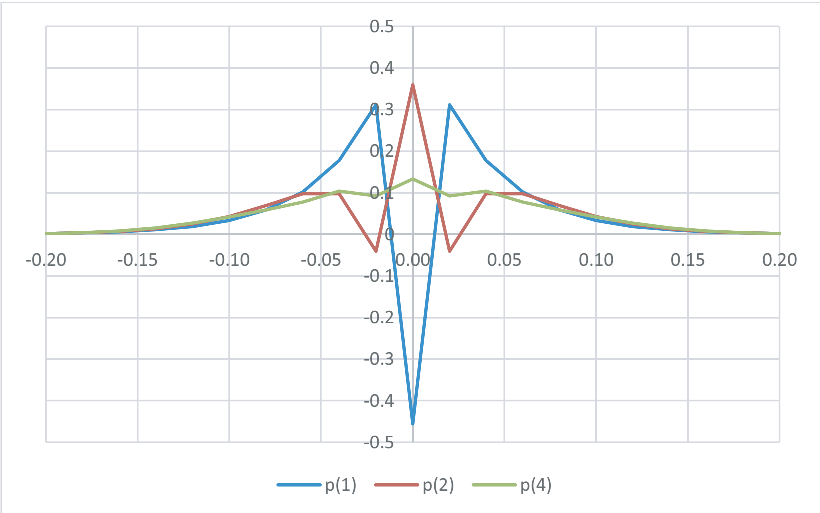
Transition probabilities from $s(0)=0$ to $s(0.5)$ with 1, 2, and 4 time steps in explicit scheme $\theta=1$. Parameters are $\mu=0, \sigma=0.1$.

Table 1: Option Prices in Finite Difference Schemes.

# steps	explicit	implicit	c-n
1	0.06250	0.02050	0.02674
2	-0.03516	0.02188	0.02321
4	-0.07683	0.02266	0.02353
closed-form	0.02349	0.02349	0.02349

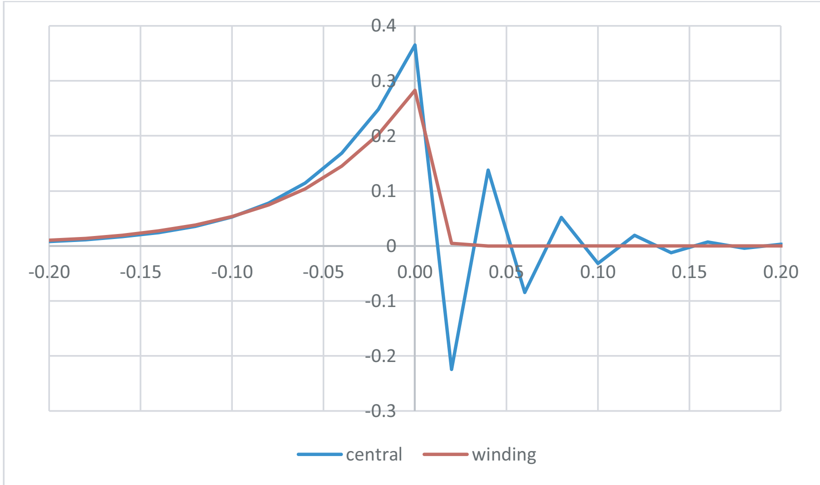
Price of option with expiry $t=0.5$) and strike $k=0.01$ in explicit, implicit and Crank-Nicolson finite difference schemes. Parameters are $s(0)=0$, $\mu=0, \sigma=0.1$.

Figure 3: Transition Probabilities for Crank-Nicolson Scheme.



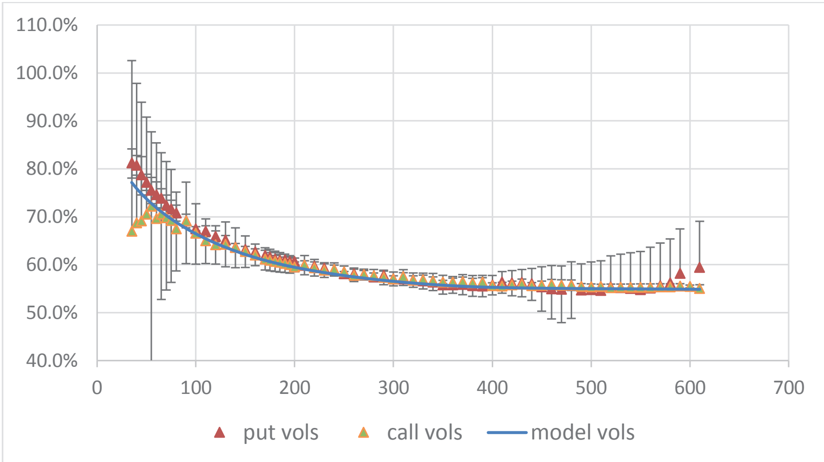
Transition probabilities from $s(0)=0$ to $s(0.5)$ with 1, 2, and 4 time steps in explicit scheme $\theta=0.5$. Parameters are $\mu=0, \sigma=0.1$.

Figure 4: Transition Probabilities — Central Differences versus Winding.



Transition probabilities for $s(0)$ to $s(0.5)$ with one time step in a fully implicit finite difference scheme, $\theta=1$, under central differencing and winding. Parameters are $\mu=-0.1, \sigma=0.01$.

Figure 5: Model versus Market



Snap shot on 8 Nov 2022 of local volatility model generated implied volatilities versus market's implied volatilities for TSLA options expiring 17 Jan 2025. Triangles show implied volatilities for mid puts and calls. Error bars show the bid offer spreads in implied volatility terms.