STRATOS
Research

Zero Days to Expiration: How they are reshaping equity markets from the inside

Market Microstructure, Gamma Risk and the New Architecture of Intraday Volatility

Fabio Iannucci

May 15, 2026

---

1. Introduction

In recent years, the options market on the S&P 500 index had a structural transformation that, despite its profound implications, has received relatively limited attention in mainstream financial discourse. At the center of this change are the so-called zero days to expiration options (0DTE): option contracts that expire the same day they are traded. What until a few years ago was a small segment has now become the most active part of the entire U.S. derivatives market.

The numbers speak by themselves. In the fourth quarter of 2024, 0DTE options on the S&P 500 surpassed all other expiration dates combined in terms of volume. The growth, therefore, has not been gradual, it has been a structural break.

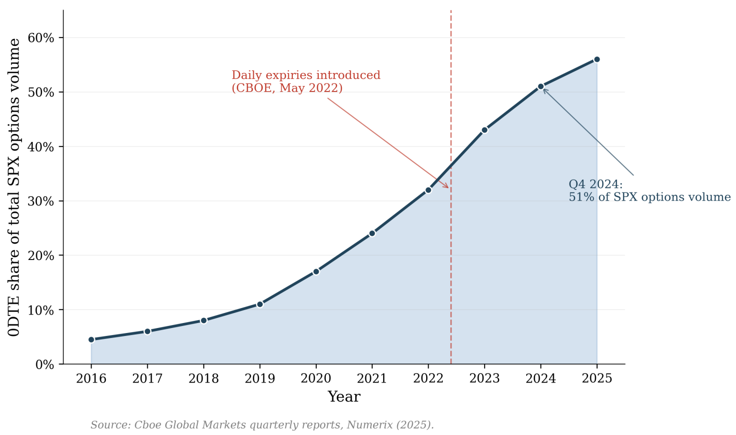

Figure 1: Growth of 0DTE options as a share of total SPX options volume, 2016–2025. The dashed line marks the introduction of daily expiries by CBOE in May 2022, after which the segment accelerated from ∼32% to over half of total volume in less than three years.

The argument developed in the rest of this paper is that the structural impact of 0DTE options does not just lie in their volume, but in the mechanical hedging response they impose on dealers, and in the way this response reshapes the microstructure of the underlying market. As time to expiration collapses, the mathematics of delta hedging breaks down, and the resulting flows transform the futures market from an instrument for price discovery into a transmission channel for options positioning, with measurable consequences on intraday volatility, on the behavior of prices in the final hours of the trading session, and on the way market makers manage their residual risk.

2. The Market Maker's Problem: Delta Hedging Under Vanishing Time

To understand why 0DTE options have such a structural impact on the underlying market, it is necessary to examine how market makers manage the risk of these positions, because this is where the mechanism begins.

Every time an operator buys a 0DTE option, on the other side of the transaction there is a market maker selling that position. The market maker is not a directional trader: its objective is to collect the bid-ask spread while keeping its book neutral with respect to market movements. To achieve this neutrality, the dealer continuously eliminates the directional risk embedded in the options it holds through *delta hedging*: it computes the aggregate delta of its position — its sensitivity to movements in the underlying — and trades an equivalent quantity of futures contracts, typically the E-mini S&P 500, to bring the net exposure to zero.

Formally, for a vanilla European option with price $V(S,t)$:

$$\Delta = \frac{\partial V}{\partial S} \tag{1}$$

where S is the price of the underlying. The dealer holds $-\Delta$ units of the underlying for each option sold, so that small movements in S leave the portfolio value approximately unchanged.

In options with standard expiries — weekly or monthly — this hedging process is relatively gradual. 0DTEs break this logic, and the reason lies in how the delta behaves as the contract approaches expiration. Under the Black-Scholes framework:

$$\Delta = N(d_1), \qquad d_1 = \frac{\ln(S/K) + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}} \tag{2}$$

When we write $\Delta = N(d_1)$, we are essentially saying that the delta of a call is a CDF of d_1, which is approximately the probability that the option finishes ITM, where d_1 measures the standardized distance from the strike. The further the underlying is above the strike, the closer this probability gets to one, and the closer the delta gets to one.

The critical term is $\sigma\sqrt{\tau}$, the effective width of the delta transition region. For a one-month option with $\sigma = 20\%$, $\sigma\sqrt{\tau} \approx 5.8\%$ and the delta transitions smoothly from zero to one over a broad range of prices. For a 0DTE with two hours to expiry, the same expression collapses to roughly 0.7%: the transition region shrinks dramatically and the delta behaves almost as a step function around the strike.

The rate at which delta changes is captured by gamma:

$$\Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{\partial \Delta}{\partial S} \tag{3}$$

For an at-the-money call, $\Gamma \propto \frac{1}{S\,\sigma\sqrt{\tau}}$, which diverges as $\tau \to 0$. Deep in- or out-of-the-money, gamma is small and delta is stable; for an at-the-money 0DTE, even a tiny move in S generates a large change in delta, forcing the dealer to rebalance its futures hedge in a very short time window.

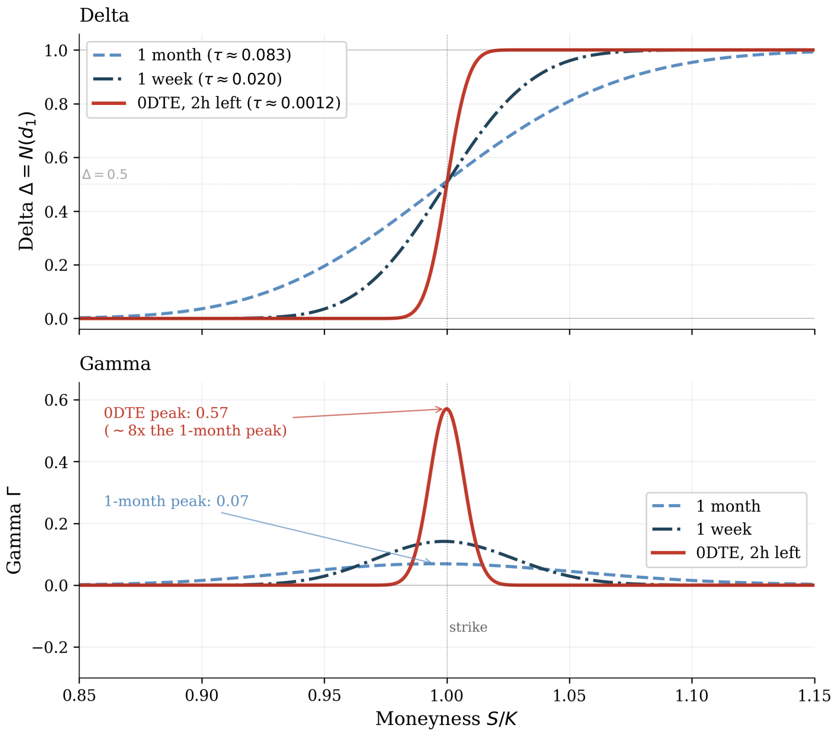

Figure 2: Delta (top) and gamma (bottom) of a European call as a function of moneyness S/K, for three times to expiration ($\sigma = 20\%$).

As shown by [3], the sensitivity of a 0DTE option to movements in the underlying cannot be adequately described by standard pricing frameworks: pure diffusion models miss the dynamics of intraday tail risks. The practical consequence is that any significant move in the S&P 500 forces dealers to rebalance massively and immediately, materializing directly in the E-mini market, compressing order book liquidity and generating price pressures that feed back into the underlying. The 0DTE segment effectively transforms the futures market from an autonomous price discovery venue into a transmission mechanism for options positioning, something that did not happen at this scale before CBOE introduced the daily expiry calendar in May 2022.

A point often missed in the public debate, raised by [4], is that gross notional volume is a misleading measure of aggregate risk. What matters is the *net imbalance* between buys and sells on each strike: if flows are perfectly balanced, the net gamma risk for the dealer is theoretically zero regardless of volume. The problem appears when this balance breaks down, generating net positions that dealers must hedge. And it is exactly in these moments of imbalance that market microstructure becomes fragile and subject to the disruptive dynamics described in the next section.

3. Short Gamma and the Amplification of Intraday Volatility

The mechanism described in the previous section is not neutral with respect to the direction of the trade. Whether the dealer needs to buy or sell futures when the market moves depends on the sign of its gamma exposure, and this distinction is at the heart of why 0DTE options can amplify rather than dampen market moves.

When a market maker is *long gamma* , that happens when it has a net long position in options , its hedging behavior is stabilizing. As the underlying rises, the delta of the long option position increases, which means the dealer must sell futures to keep the book delta-neutral. As the underlying falls, the delta decreases, and the dealer must buy. This is the classic "buy low, sell high" pattern, which absorbs liquidity shocks and tends to reduce realized volatility. The opposite happens when dealers are *short gamma*: having sold options, they must hedge the negative gamma exposure by trading in the same direction as the market. When the underlying rises, the short option position loses delta faster than a linear exposure, so the dealer must buy futures to compensate. When the underlying falls, it must sell.

More formally, consider a market maker that is short a call option with delta Δ and gamma Γ. The portfolio is hedged by holding Δ units of the underlying. Following a price move $\mathrm{d}S$, the change in the option delta is:

$$\mathrm{d}\Delta = \Gamma\,\mathrm{d}S \tag{4}$$

which means that to maintain delta-neutrality, the market maker must buy (or sell) an additional quantity $\Gamma\,\mathrm{d}S$ of the underlying. The sign of this trade goes in the same direction as $\mathrm{d}S$: positive moves trigger buying, negative moves trigger selling. This is the formal definition of *trend-following hedging*.

The aggregate effect across the market depends on the net gamma position of the liquidity providers, commonly referred to as *Gamma Exposure* (GEX). When GEX is positive, the liquidity providers are overall long gamma and the hedging flows dampen the market. When GEX is negative, the hedging flows amplify it. Following the framework developed by [1] , the realized variance of the spot market can be schematically decomposed into a baseline component and a hedging-induced component that depends on the sign and magnitude of the aggregate dealer gamma exposure:

$$\sigma^2_{\text{spot}} \;=\; \sigma^2_0 \;+\; \alpha \cdot \mathrm{GEX}_{\text{short}} \;-\; \beta \cdot \mathrm{GEX}_{\text{long}} \tag{5}$$

where σ^2_0 is the baseline variance of the underlying in absence of hedging flows, $\mathrm{GEX}_{\text{short}}$ and $\mathrm{GEX}_{\text{long}}$ are the absolute values of the aggregate short and long gamma positions, and $\alpha, \beta > 0$ are coefficients that depend on the elasticity of the futures market to order flow. The expression makes explicit that short gamma adds variance, while long gamma subtracts it.

The relevance of this framework for the 0DTE market is particularly significant. In the traditional options market, the net gamma position of dealers is relatively balanced across expiries, because large institutional investors often buy puts for protection while writing calls for yield enhancement, which puts dealers in a mixed or even long gamma state. The 0DTE segment has a different structural profile: a large part of the flow is generated by directional speculation, with retail and institutional traders buying short-dated options to take leveraged views. The liquidity providers therefore end up systematically short gamma on the 0DTE strip, and the magnitude of this short position becomes very high exactly when the gamma of each individual option is at its peak, that is, near the strike and close to expiration.

The consequence is that small price movements in the S&P 500 can generate disproportionate hedging flows in the E-mini futures, which in turn push the price further in the direction of the initial move. A move from 4500 to 4510 forces dealers to buy futures, which tends to bring the index to 4515, which forces them to buy more, and so on. The loop can self-sustain for several minutes before the liquidity providers reach a new equilibrium, or before flows from other participants absorb the pressure. This is the mechanism through which a relatively small imbalance in the 0DTE market can translate into a sharp intraday movement in the underlying, despite the absence of any fundamental news.

It is important to note that this amplification does not appear symmetrically throughout the trading day. The sensitivity of dealer hedging to price moves is itself a function of τ: as expiration approaches, gamma rises and the required hedging trades per unit of price movement become larger. The end of the day is therefore the critical window, and the phenomenon that describes this concentration of flows in the final hours of trading is the object of the next section.

4. The End-of-Day Effect

One of the most visible consequences of the 0DTE market on intraday microstructure is the modification of the volatility profile in the final hours of the trading session. What is commonly referred to as the *3:30 PM effect*, at which market quality tipically starts to deteriorate, is the empirical observation that the last hour of trading on the S&P 500 has become systematically more volatile and more prone to sharp directional moves than the rest of the session.

The 3:30 PM reference does not mean that a volume spike occurs at that exact time, but rather that the deterioration of market quality becomes observable in the thirty minutes before the close. The actual volume peak, as documented by intraday data on the E-mini futures, tipically occurs at 4:00 PM, when SPX index options settle at the closing print and all residual hedging positions must be unwound.

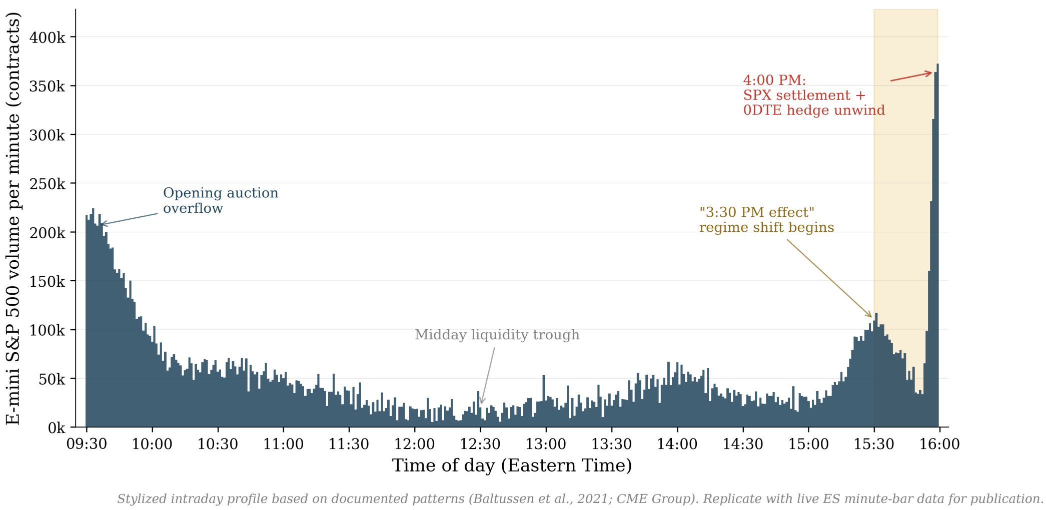

Figure 3: Stylized intraday volume profile of the E-mini S&P 500 (ES) futures .

The mechanism behind this effect is a combination of three separate events. The first is the rise in gamma described in the previous section. The second is the deterioration of liquidity in the order book itself: in the final hour of the session, many participants unrelated to options trading start to reduce their exposure ahead of the close, which means that the same hedging order from a dealer hits a thinner book and produces a larger price impact than it would earlier in the day. The third, and possibly most important, is the *unwind effect*. At 4:00 PM, the SPX options settle and all hedges must be closed: if the dealer was holding

30,000 long futures contracts to hedge a short call position that expires in-the-money, those 30,000 contracts must be liquidated within the final minutes of trading, independently from the market conditions.

The aggregate size of these end-of-day flows can be estimated from the aggregate gamma exposure of the liquidity providers. If the total short gamma position at the close is Γ_{net} and the price of the underlying has drifted by an amount ΔS during the final hour, the forced unwind volume is approximately:

$$V_{\text{unwind}} \;\approx\; \Gamma_{\text{net}} \cdot \Delta S \cdot M \tag{6}$$

where M is the contract multiplier of the E-mini futures and the expression gives an order of magnitude estimation of the number of futures contracts that dealers must trade to close their hedge before settlement.

[2] provide empirical evidence supporting this framework, showing that the hedging activity of options dealers produces a measurable momentum effect in the final thirty minutes of trading, with the sign of the effect correlated with the sign of the dealer gamma exposure. Their analysis show the explosion of the 0DTE market, but the mechanism they identify is exactly the one that the 0DTE strip has amplified: when the liquidity providers are short gamma and the underlying has been trending during the session, the unwind flows at the close reinforce the direction of the move, producing the characteristic patterns of sharp acceleration or reversal observed on many trading days.

The implications of this regime shift go beyond the technical interest of intraday microstructure. Intraday traders must now account for a structural component of volatility concentrated in the final hour, which is not driven by fundamental information but by the mechanics of the options market itself. The final hour of trading, which used to be considered a relatively neutral window for distributing executions, has now become one of the most information-poor and mechanically-driven segments of the entire session. But the participant that has been affected most directly by this transformation is not the buy-side trader: it is the market maker itself. The next section takes its perspective and looks at how the introduction of 0DTEs has reshaped the way liquidity providers think about and manage risk.

5. Risk Management Implications for Market Makers

The market maker is the agent that, in exchange for collecting the bid-ask spread, takes on the gamma risk that other participants want to offload. In the world of standard options expiries this exchange was manageable: the gamma profile of a one-month book is smooth, the hedging adjustments are gradual, and the cost of imperfect hedges can be averaged across many days. The 0DTE regime has changed the fundamental nature of this trade-off, and the way liquidity providers think about and manage their risk has had to change with it.

The first and most direct consequence is that the residual risk after delta hedging has become structurally larger. As shown in section 2, the gamma of a 0DTE option diverges as τ approaches zero, which means that the standard delta hedge fails to neutralize the position when the underlying moves through the strike. The market maker is left with a second-order exposure that cannot be fully eliminated, and that becomes particularly painful exactly in the moments when the market is most volatile. In practice, the P&L profile of a 0DTE portfolio exhibits significant discontinuity rather than scaling smoothly with trading volume. Its performance is primarily dictated by the timing and occurrence of the underlying asset crossing the strikes of peak dealer concentration.

This has had two important practical consequences. The first is operational: the frequency of hedging adjustments required by a 0DTE book is orders of magnitude higher than what a standard book demands, which has forced the major liquidity providers to invest heavily in low-latency infrastructure and automated hedging systems. Market making has always been a technology-intensive business, but in the 0DTE era the gap between the few firms with the infrastructure to operate properly and the rest of the industry has widened, with a resulting concentration of activity in a smaller number of large players. The second consequence is economic: the profitability of a 0DTE book depends almost entirely on the spread between implied and realized volatility within a single trading session. When realized exceeds implied, the gamma costs accumulated through hedging exceed the premium collected, and the position turns into a loss. The time horizon leaves no possibility to average across days or weeks, which makes the entire business significantly more volatile than traditional market making.

For the risk manager working at a market making firm, all this translates into a different set of priorities. The classical risk metrics built on historical data, such as VaR computed on multi-year time series, are poorly suited to capture the behavior of a 0DTE book, because the distributions of intraday returns under the current regime have heavier tails and a different time structure than the data observed before 2022. Standard practice has shifted toward scenario analysis based on synthetic gamma exposures, real-time monitoring of the aggregate book across strikes, and stress tests focused on the final hour of the session rather than on overnight or multi-day windows.

What emerges from this picture is that the 0DTE regime has not introduced a new type of risk, but has compressed and intensified an old one. The mathematics of options has not changed, but the speed at which it operates has, and this is enough to require a different mindset from the agents that intermediate between the buyers and sellers of these contracts. Whether this regime will remain stable or will eventually trigger the kind of stress event that some commentators have been warning about is a question that cannot be answered today. What is already clear is that the participants who manage risk on the assumption that the market still works as it did before 2022 are operating with a model that no longer matches reality.

References

[1] Benjamin Anderegg, Florian Ulmann, and Didier Sornette. The impact of option hedging on the spot market volatility. *Journal of International Money and Finance*, 122:102576, 2022.

[2] Guido Baltussen, Zhi Da, Sten Lammers, and Martin Martens. Hedging demand and market intraday momentum. *Journal of Financial Economics*, 142(1):377–403, 2021.

[3] Federico M. Bandi, Nicola Fusari, and Roberto Renò. 0DTE option pricing. Working Paper, Johns Hopkins University / ESSEC Business School, 2024.

[4] Dmitry Garmash. Zero DTE options gamma hedging. SSRN Working Paper 5329719, 2024.