Avellaneda–Stoikov: quotes, approximation and a probability above one

Igor Rivin · Deeperly / Dimension Reducers · 25 November 2026

Scoped source adjudication. This report reviews the 2008 journal article, Quantitative Finance 8(3), pp. 217–224. Two independent source-only reviews (Sol 6.2 high and Fable 5.2 high), a focused Astra 6.2 max adjudication and coordinator checks are complete. Model agreement is not a proof certificate. The retained witnesses concern specified formulas and the written simulation rule.

The formal finite-horizon utility calculation and conditional quote optimization remain useful. A separate stationary formula is incorrect, the explicit finite-horizon quotes solve an approximation, and the simulation rule needs a specified repair. These conclusions should not be collapsed into a claim that the entire market-making framework is false.

Stationary ask formula and domain

On printed p. 219, Section 2.3, define

$$D(q) = \omega - \tfrac{1}{2}\gamma^2\sigma^2 q^2.$$

Integrating the displayed frozen-inventory utility gives $\bar{v}(x,s,q) = -e^{-\gamma(x+qs)}/D(q)$, provided $D(q) > 0$. The two indifference equations therefore give

$$\bar{r}^a = s + \gamma^{-1}\log\frac{D(q)}{D(q-1)}, \qquad \bar{r}^b = s + \gamma^{-1}\log\frac{D(q+1)}{D(q)}.$$

The printed bid expression agrees with this calculation. In the printed ask expression, the denominator inside the added fraction needs $(q-1)^2$ where the paper has q^2. At $q = 0$ and $\gamma = \sigma = \omega = 1$, both required utilities are finite, but the correct ask offset is $\log 2$, whereas the printed value is $\log(3/2)$.

Both quotes require finiteness at the starting inventory and the relevant neighboring inventory. The proposed boundary choice $\omega = \gamma^2\sigma^2(q_{\max}+1)^2/2$ makes a neighboring denominator zero at an endpoint; a strict inequality or an explicit rule excluding outward quotes is needed. This isolated stationary discussion is not an input to the finite-horizon simulation formulas.

Exact control problem versus linearized quotes

Printed p. 221 explicitly introduces a linear approximation of the arrival terms. Removing the inconsistent factor $1/2$ in the q^2 term of (22) makes (23)–(30) consistent with the convention used in the ensuing calculation. Equations (29)–(30) can then be recovered within that linearized problem.

They are not an exact solution of the full nonlinear HJB. Write $\tau = T - t$, $L = \gamma^{-1}\log(1+\gamma/k)$, and $h = L + \gamma\sigma^2\tau/2$. For the frozen-inventory quadratic, the exact arrival term is

$$\frac{2A}{k+\gamma}e^{-kh}\cosh(k\gamma\sigma^2 q\tau).$$

A term independent of inventory cannot cancel this dependence for all q. Thus the earlier statement that the controlled indifference prices coincide with the frozen-inventory values requires the approximation qualifier. This does not show that an approximation is useless; an error bound and a numerical comparison with the nonlinear solution were not supplied or reproduced here.

Later work by Guéant, Lehalle and Fernandez-Tapia addresses a related problem with inventory constraints through linear ODEs and derives further approximations. Its abstract establishes this historical context; this campaign has not independently audited that later proof. It should not be presented as a repair first discovered by our panel.

The printed Bernoulli rule can fail after a reachable fill

Printed pp. 222–223 use $A = 140$, $k = 1.5$, $\sigma = 2$, $T = 1$, $dt = .005$, initially $q = 0$, with fill probabilities $p_a = Ae^{-k\delta^a}dt$ and $p_b = Ae^{-k\delta^b}dt$. The explicit quotes imply

$$\delta^a = L + (1-2q)\gamma\sigma^2(T-t)/2, \quad \delta^b = L + (1+2q)\gamma\sigma^2(T-t)/2.$$

For the Table 3 parameter $\gamma = 1$, the starting bid and ask fill probabilities are each about 0.0161973. A buy-only first step has positive probability under either ordinary independent Bernoulli draws or mutually exclusive fills. At the next step, $(q, t) = (1, .005)$, the printed ask fill probability is about **6.43715**.

The Table 1 parameter $\gamma = .1$ also fails. Three successive buy-only steps reach $(3, .015)$, where $p_a \simeq 1.16503$. Every earlier marginal probability is valid, and their sums are below one. The script records both conventional joint-fill interpretations; the original code and its treatment of simultaneous fills are unavailable.

This establishes a defect in the written simulation procedure, not falsity of the reported table entries. A cap, adaptive step or different fill rule would change the procedure and must be documented and rechecked. Using $1 - e^{-\lambda dt}$ repairs a marginal probability but, by itself, does not exactly simulate repeated fills with inventory-dependent intensities. The witness uses the displayed quotes without clipping. If negative quote distances are excluded by an intended admissibility convention, that convention instead requires changing the quoted strategy at the same state; it does not make the unmodified printed algorithm valid.

For the stated exponential intensities, shifting the quote center away from the midprice increases the sum of intensities at a fixed spread by a cosh factor. Even with the symmetric benchmark's time-averaged spread, Jensen's inequality works against the paper's explanation that centering produces more order flow. Profit and fill count are different quantities. This is a model-internal interpretation issue; no realized trade counts were available to compare.

Verification, coverage and reproducibility

The portable check bundle uses the Python standard library and 80-digit decimal evaluations. These are high-precision checks, not interval or proof-assistant certificates. Its reachable-state witnesses match the calculations retained before publication. The bundle includes the script, recorded output and instructions for a fresh run.

All three responses returned complete structured assessments. Their conservative usage estimate is **USD 2.408137**; the entire three-paper campaign used USD 8.7878365 under a USD 25 ceiling. These are usage estimates, not invoices. Publication incurred no new paid inference. Read the coverage and cost record for per-attempt details.

Full HJB admissibility and verification, the imported antecedent results, empirical order-arrival calibration and the original simulation tables remain outside this completed scope. The geometric-price appendix has not received a separate complete audit. An unstated implementation convention might explain the tables; only source code and a reproducible rerun could settle that question. No misconduct or literature-priority claim is made.

Sources and downloads

- Reviewed journal PDF.
- Author-hosted source.
- Source edition and hashes.
- This report as a PDF.
- Reproducible checks.
- Neutral reading map.

The reading map indexes all numbered equations, selected unnumbered argument units and the full bibliography. Its selected prerequisite edges are source annotations, not an exhaustive dependency proof. Panel decisions and proposed corrections are kept in the separate decorated graph. The three reviewer records retain their own hypotheses, qualifications and unaccepted proposals.

Cite this report

```
@techreport{rivin2026avellanedastoikov,
  author = {Rivin, Igor},
  title = {Avellaneda--Stoikov: Quotes, Approximation
           and a Probability Above One},
  institution = {Dimension Reducers},
  year = {2026},
  url = {https://mathreferee.dimensionreducers.ai/papers/human-avellaneda-stoikov-market-making/panel/},
  note = {AI-assisted scoped source adjudication}
}
```