

Introducción

Los mercados de apuestas son un laboratorio natural de eficiencia informacional: cada evento tiene un precio público y un desenlace verificable en horas (Levitt, 2004). La investigación avanzó en dos etapas. Primero: ¿puede un modelo con métricas sabermétricas superar al precio de cierre? Segundo, ante la respuesta negativa: si ese precio no puede superarse, ¿lo cotizan todas las casas de forma consistente entre sí? La evidencia en otros mercados apunta a que la ineficiencia explotable vive en la **dispersión entre operadores**, no en el pronóstico (Vlastakis et al., 2006).

Objetivo: cuantificar el techo informacional de los modelos predictivos con datos públicos y validar una fuente de ventaja que no dependa de predecir mejor que el mercado.

¿Por qué es difícil?

- El rival es un pronóstico agregado. El precio de cierre incorpora a miles de participantes con dinero en riesgo.
- Hay un peaje. El margen de la casa (~4.5%) debe superarse antes de ganar un solo peso.
- Poca señal. El mejor equipo pierde 4 de cada 10 juegos: el techo predictivo está a pocos puntos del azar.
- El riesgo metodológico domina al estadístico. Una fuga mínima fabricó +8.4 puntos ficticios.

Metodología

- Datos:** 6,840 partidos MLB 2023–2026; 201,112 registros de bateo y 75,276 de pitcheo; Statcast, Bayes empírico, clima, *umpires* y *lineups*. 66 variables.
- Etap 1 — Predicción.** Ensamble `glmnet` + `nnet` + `bayesglm` con validación *walk-forward* y auditoría *as-of* estricta: cada variable se reconstruye con la información disponible antes del primer lanzamiento.
- Etap 2 — Microestructura.** Momios de apertura y cierre de seis casas, MLB 2021–2026. A cada casa se le remueve el margen; el consenso de las demás estima el precio justo (Shin, 1993; Štrumbelj, 2014); si el mejor precio paga más que ese consenso, hay valor.
- Regla congelada** el 24-jul-2026 con SOLO 2023–24 y aplicada sin cambios al resto: umbral 2%, mínimo 3 casas en el consenso. Dimensionamiento por Kelly fraccional con control de CVaR (Rockafellar & Uryasev, 2000).
- Juez:** CLV = precio tomado × probabilidad justa de cierre – 1, con *devig* de Shin y prueba t unilateral. ROI como KPI secundario.
- Etap 3 — Validación en vivo.** Cuatro hipótesis *pre-registradas* el 24-jul-2026, antes de observar un solo dato, con captura propia multi-liga desde el 25-jul-2026 y *kill-switch* declarado.

Resultados

1 El modelo no supera al mercado

La precisión aparente de 60.55% en *walk-forward* estaba inflada por fuga de información (*look-ahead* en variables de temporada). La precisión honesta *as-of* fue **52.11%**, indistinguible del precio de cierre. Es un resultado negativo, pero delimita el **techo informacional de los datos públicos**: diez frentes se probaron con *kill-switch* pre-especificado y nueve murieron.

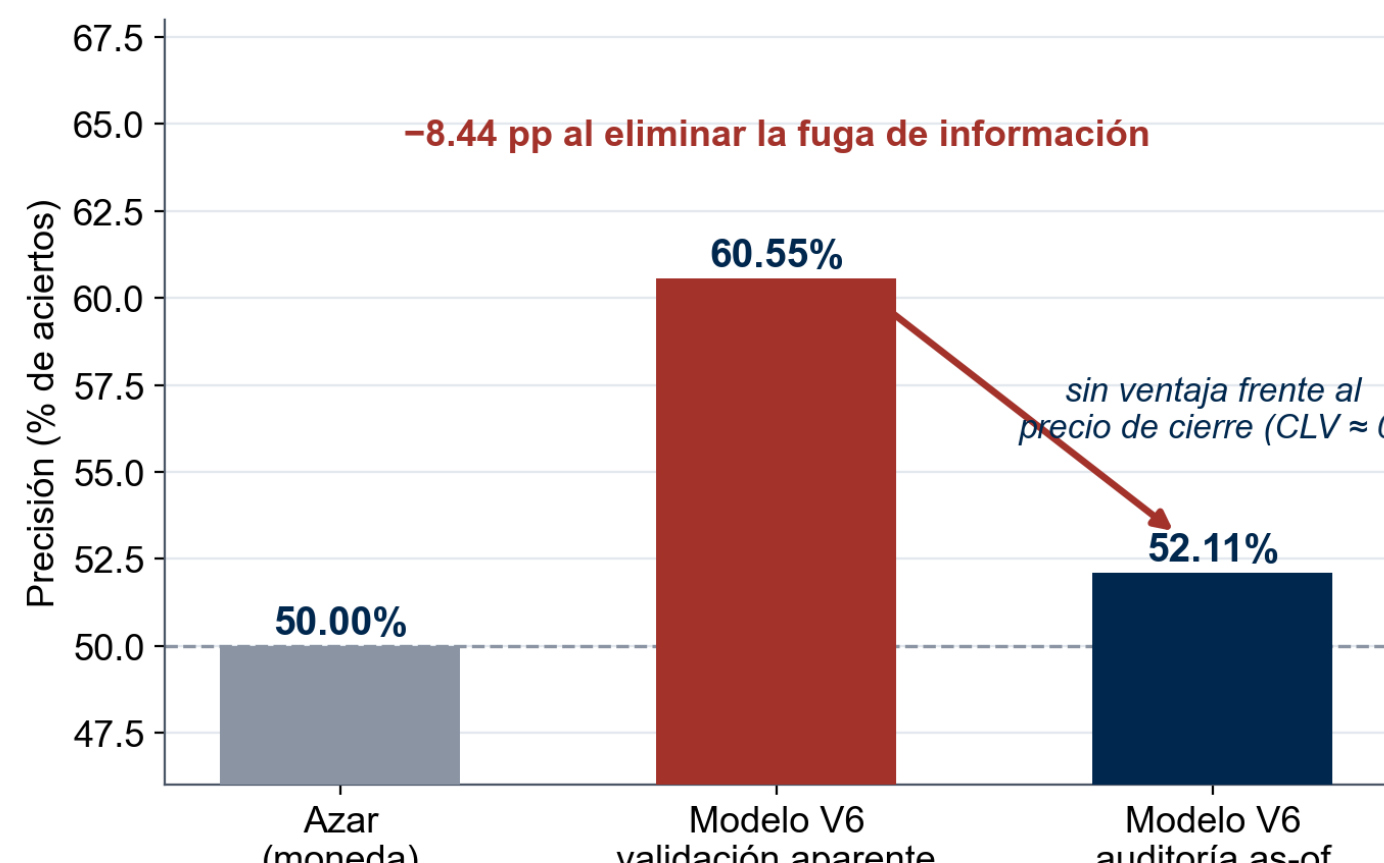


Figura 1. El costo de auditar: la ventaja aparente desaparece al reconstruir cada variable *as-of*. Fuente: elaboración propia.



Figura 2. Diez frentes probados con *kill-switch* pre-especificados; uno sobrevive. Fuente: elaboración propia.

2 Sí existe ineficiencia ENTRE casas

Temporada	Rol	n	ROI	p (ROI)	CLV (Shin)	p (CLV)
2021	nunca usada	468	+5.42%	0.21	+8.10%	1.1×10 ⁻³
2022	nunca usada	503	-1.37%	0.60	+3.41%	4.1×10 ⁻¹³
2025	holdout	202	+1.24%	0.44	+2.11%	2.2×10 ⁻³
2026	simulación	117	-5.58%	0.71	+4.30%	7.2×10 ⁻⁶
Pool	—	1,290	+1.12%	0.38	+4.99%	2.2×10 ⁻⁷

Tabla 1. Desempeño *out-of-sample* de la regla congelada, cuatro temporadas. Fuente: elaboración propia.

3 Mecanismo: ¿por qué sobrevive la dispersión?

La ineficiencia no exige que alguna casa se equivoque en promedio, sino que **no todas se actualizan al mismo tiempo**. En el panel mexicano capturado con marca de tiempo propia, el *payout* del mismo evento varía entre 90.6% y 97.5% —dispersión de hasta **27%**, un orden de magnitud mayor que en el panel estadounidense— y la casa más rápida mueve precio cada 35 minutos mientras la más lenta tarda más de siete horas. Ninguna actúa como líder de precio, de modo que la ventaja no viene de saber más que el mercado, sino de **comprar en el rezagado el precio que el consenso ya corrigió**.

CLV +4.99% · p = 2.2×10⁻⁷ · n = 1,290

Positivo en 4 de 4 temporadas *out-of-sample*, incluidas dos nunca usadas para diseñar la regla

Lo que se descartó

4 Hallazgo metodológico y frentes cerrados

- El CLV hereda el sesgo del *devig*. El método proporcional infla la probabilidad justa de los *longshots* y fabrica CLV espurio (t = +11.7 con ROI de -10% a -15% en 11,668 partidos de fútbol). Shin lo desinfla: +1.9% → +0.2%. **Regla derivada: todo CLV se mide con Shin**, y el resultado principal sobrevive al cambio de juez.
- Fútbol europeo entre casas: ROI negativo en las 15 celdas del *grid* (-9% a -58%). KILL.
- NFL en casas mexicanas: ~2 *picks* por temporada, KILL por volumen. **Props de ponches modelados**: Brier peor que el mercado (n = 1,868), KILL.
- Predicción del movimiento de línea: AUC 0.49 — el mercado no anticipa su propio movimiento; lo aprovechable es la *latencia* entre casas, no la anticipación.

5 La dificultad de demostrar una ventaja: el dinero no alcanza como juez

El ROI agrupado (+1.12%) no alcanza significancia ni con 1,290 apuestas. Con una desviación estándar de 1.26 unidades por apuesta y una ventaja de 0.011, detectar el *edge* por la vía monetaria exigiría del orden de **34,000 apuestas** —más de veinte temporadas—, mientras que el CLV lo detecta con mil. **Esa imposibilidad práctica es en sí misma un resultado** y justifica decidir por CLV. La simulación 2026 lo ilustra: proceso correcto y resultado monetario negativo conviven sin contradicción. Se reporta deliberadamente la curva perdedora, porque publicar solo la curva —ganadora o perdedora— es publicar varianza, no proceso.

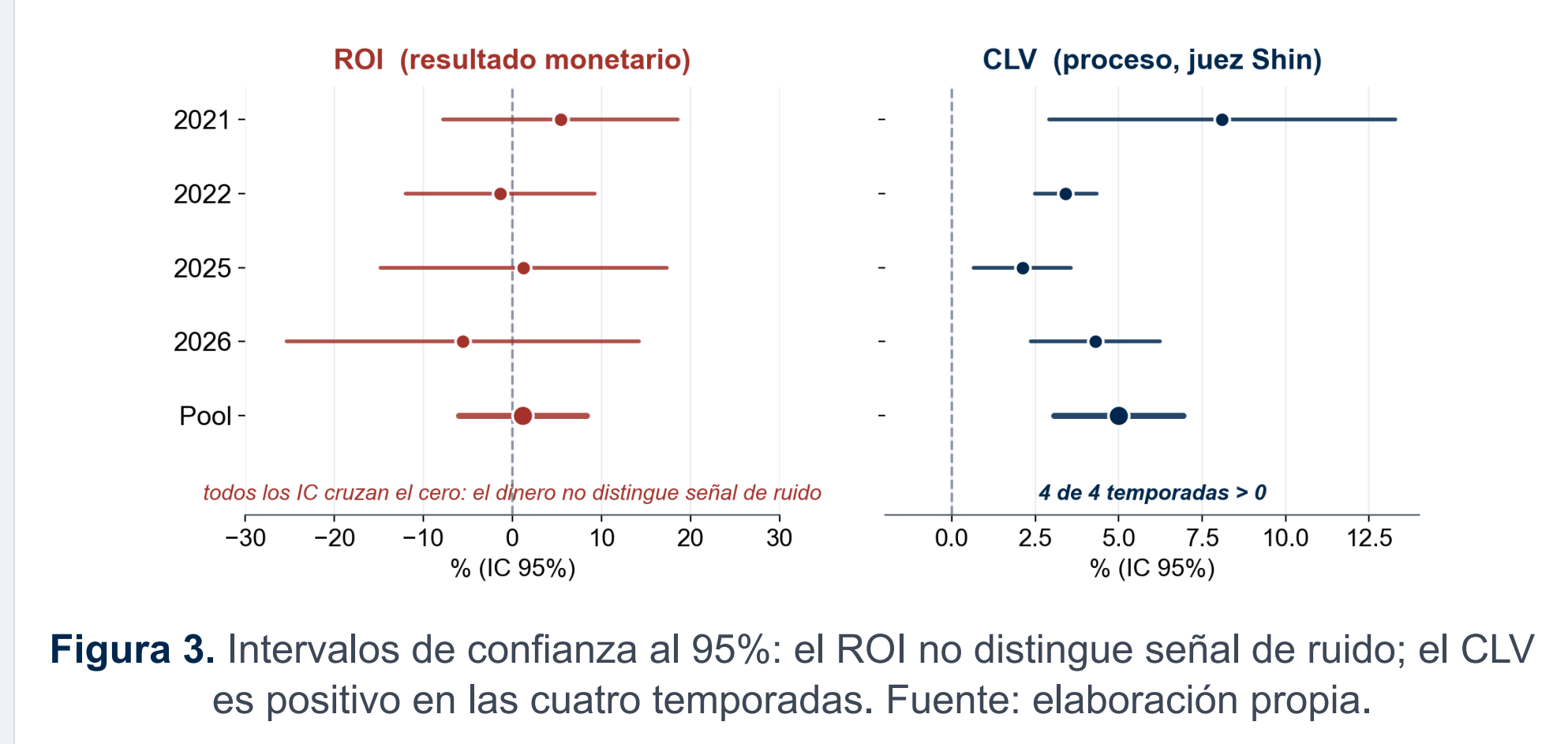


Figura 3. Intervalos de confianza al 95%: el ROI no distingue señal de ruido; el CLV es positivo en las cuatro temporadas. Fuente: elaboración propia.

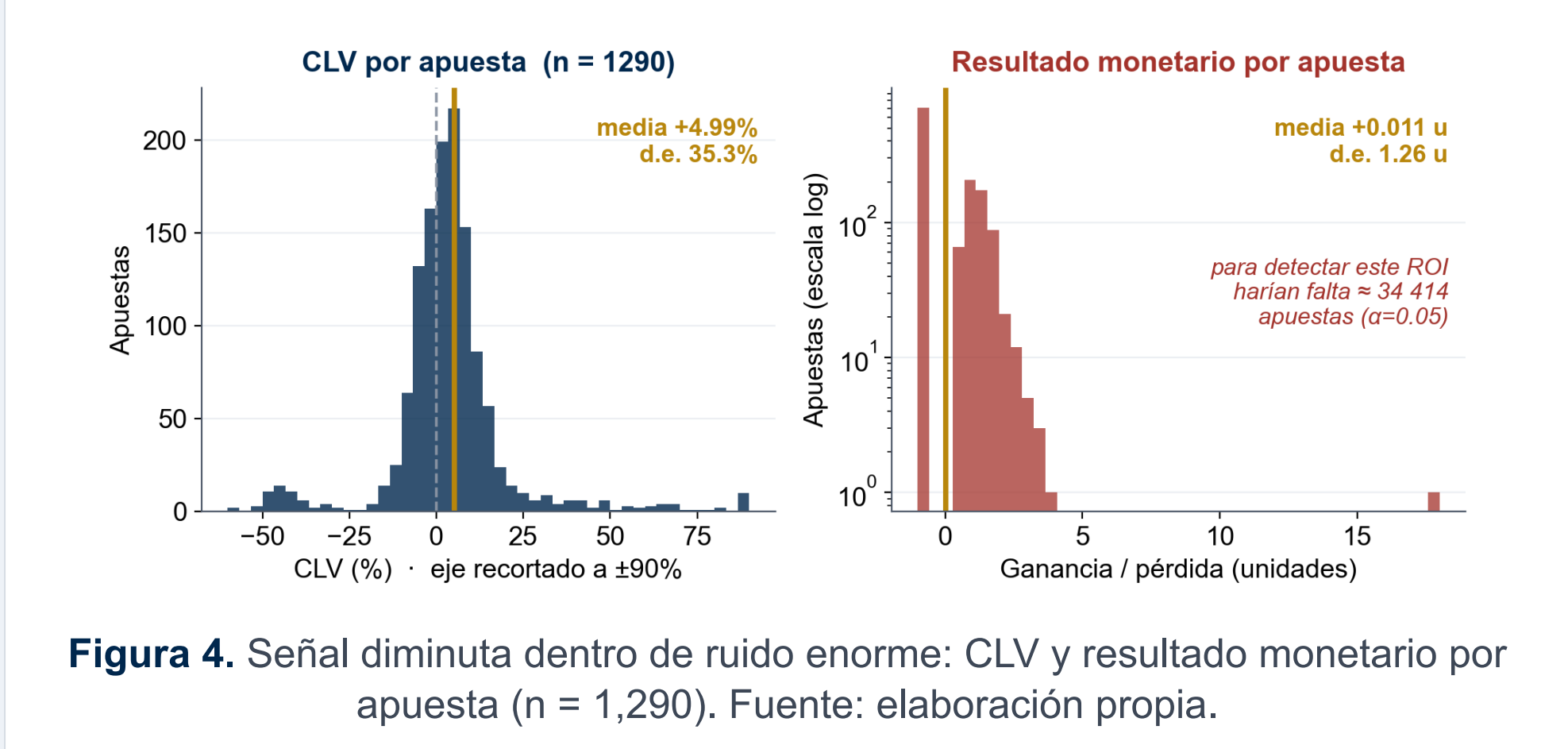


Figura 4. Señal diminuta dentro de ruido enorme: CLV y resultado monetario por apuesta (n = 1,290). Fuente: elaboración propia.

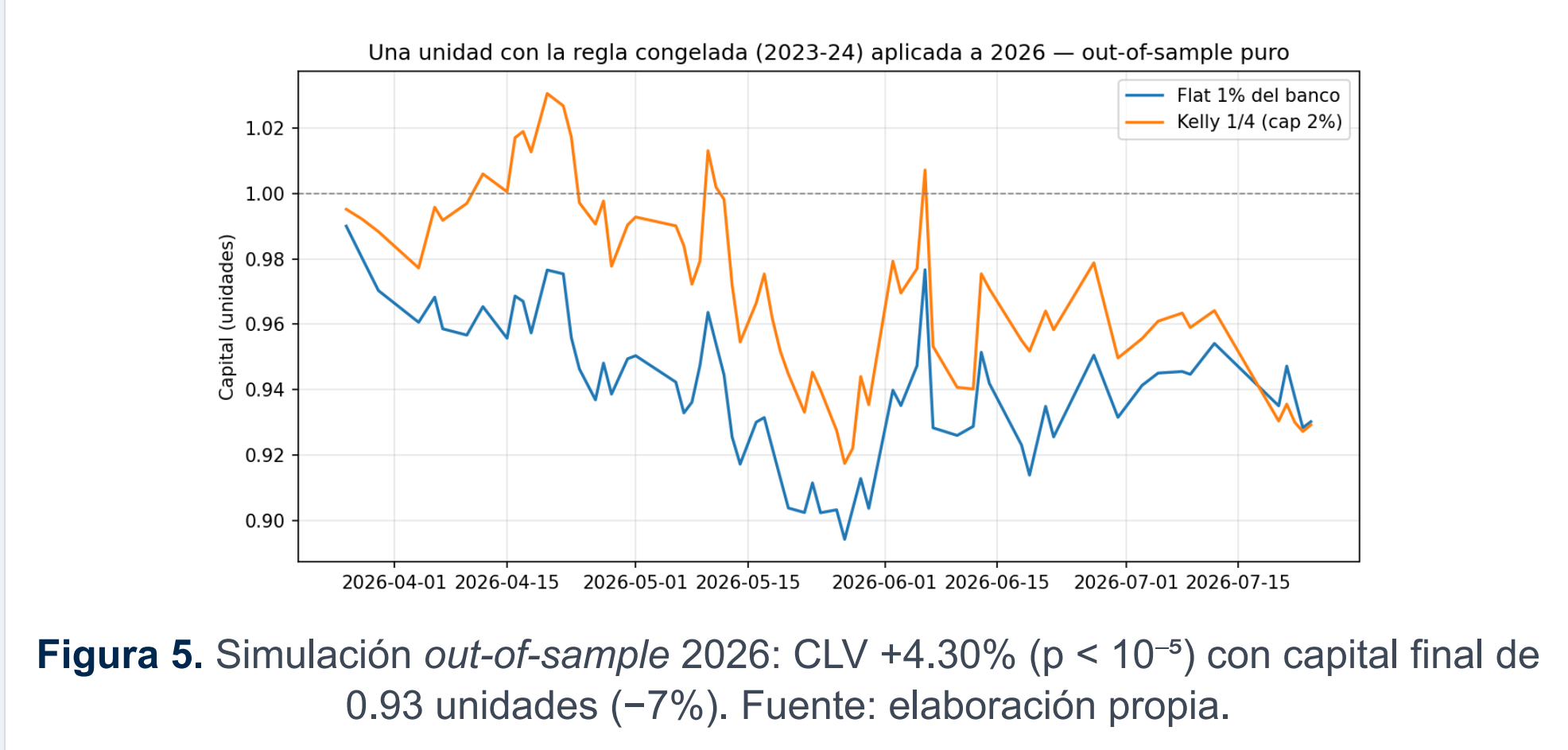


Figura 5. Simulación *out-of-sample* 2026: CLV +4.30% (p < 10⁻⁵) con capital final de 0.93 unidades (-7%). Fuente: elaboración propia.

Al mercado no se le gana prediciendo mejor: se le gana **comprando más barato** en la casa que aún no actualiza su precio.

52.11%

Precisión honesta *as-of* del modelo:
empata al precio de cierre

+4.99%

CLV *out-of-sample* entre casas
en 1,290 apuestas

2.2×10⁻⁷

Valor p del CLV agrupado:
positivo en 4 de 4 temporadas

CONCLUSIONES

- Con información pública, la MLB es eficiente en sentido semi-fuerte: el modelo **empata** al precio de cierre (52.11% *as-of*).
- La ineficiencia no está en la predicción sino en la **microestructura**: la dispersión entre casas produce CLV +4.99% (p = 2.2×10⁻⁷) en 1,290 apuestas *out-of-sample*.
- La disciplina de *kill-switch* y la elección del juez (CLV con Shin) importan más que la complejidad del modelo.
- La validación en vivo pre-registrada decidirá si el *edge* es ejecutable con costos y límites reales.

Referencias

Levitt, S. D. (2004). Why are gambling markets organised so differently from financial markets? *The Economic Journal*, 114(495), 223–246.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41.

Shin, H. S. (1993). Measuring the incidence of insider trading in a market for state-contingent claims. *The Economic Journal*, 103(420), 1141–1153.

Štrumbelj, E. (2014). On determining probability forecasts from betting odds. *International Journal of Forecasting*, 30(4), 934–943.

Vlastakis, N., Dotsis, G., & Markellos, R. N. (2006). How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, 28(5), 426–444.