

Carry Incentives and Net Warrant Accumulation: Evidence from SHFE Aluminum

Yuchen Han

F24092100015@cityu.edu.mo

City University of Macau

Research Article

Keywords: Commodity futures, Carry incentives, Registered warrants, Net inventory changes, SHFE aluminum

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11232379/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Carry Incentives and Net Warrant Accumulation: Evidence from SHFE Aluminum

Yuchen Han^{1*}

^{1*}City University of Macau, Macau, China.

Corresponding author(s). E-mail(s): F24092100015@cityu.edu.mo;

Abstract

Exchange-registered inventories may respond to the incentives to carry commodities into future delivery. We ask whether these incentives are associated with subsequent net warrant accumulation in Shanghai Futures Exchange (SHFE) aluminum. Using daily data from 2015 to 2025, we compare net warrant changes following fitted futures slopes above and below a financing-and-storage benchmark. In the unadjusted comparison, net changes are approximately 1,380 tonnes higher following above-benchmark sessions. The association remains positive after allowing for recent warrant changes, but statistical precision weakens when recent changes, year effects, and delivery-calendar controls enter together. Local benchmark and crossing-event tests also yield imprecise estimates. Daily warrant totals measure the balance of registrations and cancellations; they do not reveal gross registrations or physical warehouse flows. The evidence clarifies how carry conditions relate to net warrant accumulation in a metal delivery system, while distinguishing the broad market-state association from a discrete registration threshold. Omitted costs and common market conditions limit causal and arbitrage interpretations.

Keywords: Commodity futures; Carry incentives; Registered warrants; Net inventory changes; SHFE aluminum

JEL Classification: G13; G12; Q02

1 Introduction

Exchange-registered inventories connect a commodity's physical stocks to its futures delivery system. A warehouse may hold metal that is not registered for futures delivery, while outstanding warrants represent a particular legal claim on approved metal. The decision to maintain that claim can vary with the return to carrying metal, existing delivery obligations, and inventory management. This raises a question about inventory adjustment: when a futures curve offers stronger carrying incentives, does the outstanding stock of registered warrants subsequently accumulate more rapidly?

The theory of storage provides the economic starting point. Intertemporal price differences reflect carrying costs and the benefits of holding inventories (Working 1949; Brennan 1958; Fama and French 1987). A sufficiently steep futures curve may make carrying metal more attractive to some holders. Applying this logic to registered warrants requires care, however. Price differences can be measured in several ways, financing costs differ across holders, and the observed change in warrants combines additions with cancellations. The economically relevant empirical question is therefore whether carrying conditions are associated with the net change in a delivery-eligible stock. A directional association alone does not identify the holder's transaction or its profitability.

Existing research establishes that the definition and accessibility of inventory matter for commodity prices. Geman and Smith (2013) examine storage relationships across London Metal Exchange (LME) base metals and incorporate Chinese exchange inventories. Geman and Scheiber (2017) study SHFE copper curves and compare exchange inventories with stocks augmented by bonded metal. Kim et al. (2017) distinguish inventories in China, nearby Asian locations, and the wider LME network when examining convenience yields and imports. These studies motivate attention to the particular inventory object observed. Their stock-price relationships do not, by themselves, identify the gross flows underlying changes in registered warrants.

A complementary literature studies inventory adjustment and delivery instruments. Uría and Williams (2007) relate daily natural-gas storage flows to futures spreads, local price differences, and operating cycles. Joseph et al. (2016) match grain spreads with deliverable stocks and distinguish physical inventory from shipping certificates. Garcia et al. (2015) show why the storage charge on a delivery instrument can matter for futures convergence. Most closely, Fernandes et al. (2025) analyze corn deliveries, their timing, new shipping certificates, and feedback into futures prices. Carry incentives and delivery choices are thus established research subjects. The remaining question addressed here concerns daily *net* accumulation of SHFE aluminum warrants, whose institutional meaning and observed outcome differ from corn delivery intentions, new certificate registrations, and repeated deliveries.

Our contribution is a state comparison linking the SHFE aluminum curve to the subsequent net change in registered warrants. Using contract-level settlements, daily warrant data, and a public short-term financing benchmark, we compare the subsequent flow across carry conditions. The design separates a broad above-benchmark state from local behavior around its dividing point. That distinction matters because a positive state difference can arise without a sharp response at a particular cost estimate. Weekly exchange stocks help describe the inventory setting, but the main estimand is a daily flow and does not condition on same-frequency total warehouse stocks.

From January 2015 to April 2025, average net warrant changes are about 1,380 tonnes higher after above-benchmark sessions. The sign remains positive in specifications including recent net changes and in comparisons excluding the large 2017–2018 inventory expansion. Precision weakens when recent changes, year effects, and delivery-calendar controls are combined. Local piecewise regressions and crossing-event comparisons also provide limited evidence for a discrete response at the benchmark. The findings describe net accumulation under different carry conditions. They leave gross registrations, cancellations, executable arbitrage returns, and causal direction unresolved.

2 Institutional setting and data

SHFE aluminum is physically delivered through standard warehouse warrants. The exchange’s 2019 manual describes the aluminum contract and the transfer of valid warrants in delivery (Shanghai Futures Exchange 2019). Registration and physical storage are related but distinct: approved warehouses hold metal, while outstanding warrants record claims that can satisfy delivery requirements. A change in ownership through delivery need not change the total number of warrants, and cancellation of a warrant need not establish when the underlying metal physically leaves storage. These institutional distinctions determine what the public aggregate data can measure.

The daily panel covers 5 January 2015–25 April 2025 and contains 2,506 trading sessions. Contract-level settlements and daily aluminum warrant totals were retrieved from SHFE through AKShare; 496 original weekly warehouse reports are also available

(Shanghai Futures Exchange 2025). Daily totals are matched by exchange date, with no interpolation of missing warrant observations. The completed sequence includes two daily source records absent from the earlier cached panel. Consequently, first differences span consecutive observed exchange sessions and yield 2,505 usable outcomes. The earlier incomplete sample is retained only as a version comparison in Online Resource 1.

Let W_t denote registered aluminum warrants, measured in tonnes, on trading session t . The principal outcome is

$$\Delta W_t = W_t - W_{t-1}. \quad (1)$$

It measures net accumulation: new registrations less cancellations, potentially including revisions to reported holdings. The daily aggregate does not separately identify these components. We therefore use “net warrant change” throughout. Neither positive changes nor negative changes can be read directly as gross registrations, delivery volume, or physical warehouse inflows and outflows.

Weekly visible stocks S are the report’s spot-weight field, and weekly registered warrants are its registered-futures-weight field. Aluminum is selected by the exact product identifier, or the exact aluminum name in older reports, and its own grand-total row. This avoids mixing aluminum with alumina. The corrected mean values are 343,956 tonnes for S and 217,332 tonnes for W ; their correlation is 0.971. The median registered share is 0.523, with a range from 0.024 to 0.950. Weekly and daily warrants agree exactly on 495 of 496 common dates and within one tonne on all 496. The remaining one-tonne difference is present in the original reports and is retained explicitly. Identifiable bonded rows in the weekly aluminum reports have zero registered weight; this finding does not establish the complete tax-status composition of every daily aggregate.

For each day, the nearby price P_t is the settlement of the shortest remaining-maturity contract with positive open interest. The curve sample uses other aluminum contracts with positive open interest and one to twelve months of remaining maturity. Remaining maturity $M_{i,t}$ is the integer difference between the delivery month and the exchange-date

month. The fitted slope comes from

$$F_{i,t} - P_t = a_t + b_t M_{i,t} + u_{i,t}, \quad (2)$$

where $F_{i,t}$ is contract i 's settlement. Each day has ten or eleven qualifying contracts; at least four contracts and three distinct maturities are required. The unit of b_t is Chinese yuan (CNY) per tonne per month. Settlements summarize the exchange date and do not supply intraday executable quotes. Subtracting a common reference price leaves an ordinary least-squares (OLS) slope unchanged when the contract set and intercept are held fixed.

The financing-and-storage benchmark is

$$c_t = P_t \frac{r_t}{12} + s \frac{365.25}{12}, \quad (3)$$

where r_t is the one-month Shanghai Interbank Offered Rate (SHIBOR) expressed as an annual decimal and s is storage in CNY per tonne per day. The SHIBOR series was obtained from Eastmoney's redistribution of the public benchmark (Eastmoney 2026), and all sample sessions have an exact-date quote. The maintained storage assumption is $s = 0.45$, with 0.40 and 0.60 scenarios. An official 2019 schedule lists 0.40 for duty-paid aluminum yard storage and separate bonded charges; the 2026 snapshot provides further institutional context (Shanghai Futures Exchange 2019, 2026). These records supply dated fee evidence, not a complete warehouse-specific tariff history for the sample.

Equation (3) is a cost proxy beside a statistical curve slope. The nearby futures settlement is not an observed spot acquisition price, and SHIBOR is not every holder's financing rate. Financing tenors are not matched separately to each deferred contract. Margin funding, delivery and inspection charges, tax and invoicing conditions, transaction costs, and execution constraints are omitted. Such omissions can change the state classification; their effect on the estimated association is not known. We accordingly avoid interpreting $b_t > c_t$ as a measured violation of an executable arbitrage bound.

3 Empirical design

Define the gap $g_t = b_t - c_t$ and the state $D_t = \mathbf{1}\{g_t > 0\}$. The unadjusted comparison is

$$\Delta W_t = \alpha + \gamma D_{t-1} + e_t. \quad (4)$$

The coefficient γ is a difference in mean net changes between sessions following the two states. It is not the mean increase within the above-benchmark group. The one-session lag specifies the order of the comparison and avoids a simultaneous same-date regression. It does not establish when all information was available intraday, rule out anticipation, or identify a causal effect of the curve on warrants.

The designated primary adjusted specification adds recent net warrant changes and common calendar conditions:

$$\Delta W_t = \alpha + \gamma D_{t-1} + \rho \Delta W_{t-1} + \eta_{y(t)} + \mu_{m(t)} + \lambda_{h(t)} + e_t. \quad (5)$$

Here $\eta_{y(t)}$ is a year fixed effect, $\mu_{m(t)}$ is a calendar-month fixed effect, and $\lambda_{h(t)}$ is a fixed effect for calendar days remaining to the nearby contract's last trading date. Recent net changes capture short-run continuation in the flow. Year effects allow market conditions to differ across the inventory cycle, while calendar controls allow systematic seasonality and delivery timing. These controls enter together in the primary adjusted model. Separate-control specifications show how estimates change but do not substitute for that joint comparison.

The calendar starts from the contract-month 15th or the next exchange session. The historical manual also permits separate Spring Festival adjustments (Shanghai Futures Exchange 2019). Comparing this rule with each contract's final observed exchange quotation identifies three covered exceptions: February 2015, February 2018, and February 2021. The revised calendar uses 10 February 2015, 9 February 2018, and 5 February 2021, respectively. Selected original exchange market files confirm the contracts' presence on these dates and absence on the following session. The other completed nearby contract

months match the general rule in the cached quote history. The final nearby contract expires beyond the sample and retains the generic May 2025 date. This construction is documented separately from a claim to have recovered every historical notice.

Inference clusters observations by calendar year. We report finite-sample-adjusted cluster-robust (CRV1) standard errors and two-sided restricted wild cluster bootstrap- t p-values using Webb’s six-point weights (Cameron et al. 2008; Webb 2023). The random seed is retained from the earlier analysis, and all reported tests use 9,999 draws. The implementation records valid and degenerate draws rather than counting degeneracy as a non-rejection. Standard errors and bootstrap p-values are distinct summaries. Descriptive figure intervals use a stated $t(G - 1)$ approximation, where G is the number of contributing years; they are not bootstrap-inverted intervals.

There are only eleven year clusters, and the above-benchmark observations are concentrated in a few years. Wild bootstrap inference is useful under limited cluster support, but its availability does not make that support sufficient (MacKinnon and Webb 2017, 2018). Our state varies within years, unlike a design assigning treatment to entire clusters. Clustering also allows within-year dependence while leaving possible dependence across year boundaries untreated. It does not repair measurement error, common market shocks, or model misspecification.

To distinguish a broad state association from behavior at the dividing point, we estimate

$$\Delta W_t = \alpha + \beta g_{t-1} + \theta D_{t-1} + \kappa g_{t-1} D_{t-1} + e_t. \quad (6)$$

The jump term θ and change-in-slope term κ are reported for the full sample and fixed windows of ± 10 , ± 20 , and ± 40 CNY per tonne per month. These are local association diagnostics. The benchmark is researcher constructed and crossing it is not an externally assigned treatment, so the regressions do not provide a causal regression-discontinuity design.

A complementary event comparison defines a crossing as the first above-benchmark signal after five consecutive sessions at or below the benchmark. The response is the sum

of the following five sessions' net changes, beginning one session after the signal. Only complete response horizons enter. We also report a continuous-gap model, storage scenarios, liquidity-based curve alternatives, year exclusions, and reverse predictive regressions. Their purpose is to describe sensitivity and limits of the selected comparison.

4 Results and limitations

4.1 The state comparison and its support

The fitted slope averages 7.4 CNY per tonne per month, while the benchmark averages 48.8. One-month SHIBOR averages 2.72 percent over the sample. Figure 1 shows that slopes usually lie below the benchmark and that above-benchmark sessions occur in distinct stretches. The resulting state is a feature of the observed curve and maintained cost parameters. It should not be read as a daily observation of a holder's net trading profit.

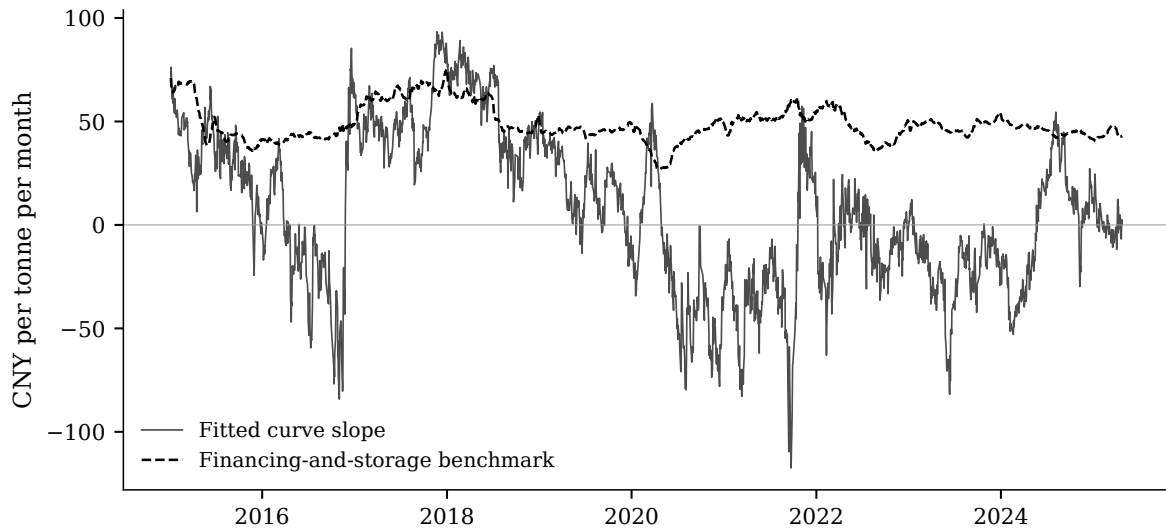


Fig. 1 Fitted curve slope and the financing-and-storage benchmark

Notes. SHFE aluminum, 5 January 2015–25 April 2025. The solid line is the daily OLS slope from equation (2). The dashed line is equation (3), using same-date one-month SHIBOR and the maintained storage assumption of 0.45 CNY per tonne per day. Both series are CNY per tonne per month. Source: SHFE contract settlements and the redistributed SHIBOR series.

Table 1 records the support for the lagged state. There are 300 above-benchmark observations, approximately 12 percent of the outcome sample. Of these, 186, or 62 percent, occur in 2017–2018. Eight years contain such observations, but their contributions are uneven: 2021 supplies only two and three years supply none. Counting

years with at least one state observation therefore overstates how evenly the empirical information is distributed. The concentration makes year exclusions informative and limits interpretation of precision from the pooled comparison.

Table 1 Sample support for the lagged above-benchmark state

Year	Observations	Above-benchmark	Within-year share (%)
2015	243	38	15.6
2016	244	17	7.0
2017	244	57	23.4
2018	243	129	53.1
2019	244	10	4.1
2020	243	30	12.3
2021	243	2	0.8
2022	242	0	0.0
2023	242	0	0.0
2024	242	17	7.0
2025	75	0	0.0
Total	2,505	300	12.0

Notes. The outcome sample is 6 January 2015–25 April 2025. Counts refer to the previous trading session’s state, D_{t-1} . The 2025 sample is partial. The completed daily panel contains 2,506 sessions; the first is lost when computing the net change. There are 37 above-benchmark spells and 19 crossings preceded by five below-benchmark sessions.

The conditional means clarify the economic magnitude. Following the above-benchmark state, the mean net change is +1,223 tonnes over 300 observations. Following the other state, it is -156 tonnes over 2,205 observations. Their difference is 1,380 tonnes. The benchmark coefficient thus describes higher *net changes relative to another group*; it does not imply that warrants increase by that amount whenever the state occurs. Individual positive-state sessions can also show large decreases, and the mean cannot identify how much gross registration or cancellation occurred.

4.2 Joint controls and economic interpretation

Table 2 puts the unadjusted state comparison beside the adjusted specifications and the linear-gap alternative. The unadjusted coefficient has a year-clustered standard error of 432 tonnes and a Webb p-value of 0.008. Including the previous session’s net change reduces the estimate to 756 tonnes. Recent changes and the state therefore contain overlapping predictive information about the subsequent flow. This attenuation is an

empirical change in the conditional comparison, not a decomposition of the baseline into independently identified mechanisms.

Table 2 Net warrant accumulation across lagged carry conditions

	(1)	(2)	(3)	(4)	(5)
Lagged above-benchmark state	1,380 (432)	756 (283)	1,200 (447)	818 (328)	
Lagged continuous gap					6.53 (8.06)
Webb bootstrap p	0.008	0.010	0.097	0.091	0.560
Lagged net change	No	Yes	No	Yes	No
Year fixed effects	No	No	Yes	Yes	No
Delivery-calendar fixed effects	No	No	Yes	Yes	No
Observations	2,505	2,504	2,505	2,504	2,505

Notes. The outcome is ΔW_t , tonnes per trading session. State coefficients are differences in net changes across states; the continuous gap coefficient is tonnes per CNY/tonne/month. Column (1) is the unadjusted comparison. Column (4) is the designated primary adjusted specification. Delivery-calendar controls are calendar-month and calendar-days-to-last-trade fixed effects, with three historical Spring Festival adjustments. Parentheses contain standard errors. All columns have 11 year clusters. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

Year and corrected delivery-calendar controls together give a positive difference of 1,200 tonnes, with bootstrap p-value 0.097. Adding recent net changes to those controls yields 818 tonnes in the primary adjusted specification, with p-value 0.091. Both estimates are less precise than the unadjusted comparison. The table accordingly supports a positive conditional association but provides weaker conventional significance under the full control set. Separate year or calendar controls cannot justify describing the result as fully robust to all necessary controls. Common market conditions that remain outside the model may still influence both the curve and warrant accumulation.

The linear-gap specification gives 6.53 tonnes per unit of gap, with bootstrap p-value 0.560. This estimate is imprecise. The state model and continuous model summarize different features of the data, but their different p-values do not themselves test which functional form is correct. In particular, a significant broad state comparison and an imprecise linear coefficient do not establish a discontinuity at zero. The benchmark's measurement uncertainty also limits the economic meaning of a sharp numerical cutoff.

The defined sensitivity comparisons appear in Online Resource 1, Section B. Excluding 2017–2018 leaves a positive difference of 1,950 tonnes. Volume weighting gives 2,045 tonnes, and restricting the curve to traded contracts in the first six months also gives a positive difference. These alternatives address liquidity and state concentration; they do not price an executable trade. In the longer maturities, some settlements occur on zero-volume contract-days, so a statistical curve can contain prices without same-day transactions. The alternative slopes retain missing values before classification, avoiding the treatment of a missing slope as an observed below-benchmark state.

4.3 Local behavior and crossing events

Figure 2 displays mean next-session net changes across fixed bins of the lagged gap. The pattern is consistent with higher average accumulation in some above-benchmark bins. However, several adjacent means are imprecise and draw on fewer contributing years than the full sample. Each interval concerns its own conditional mean; whether an interval excludes zero does not test equality with the neighboring mean. Binning is therefore a descriptive aid to the local regressions, not an independent demonstration of a threshold.

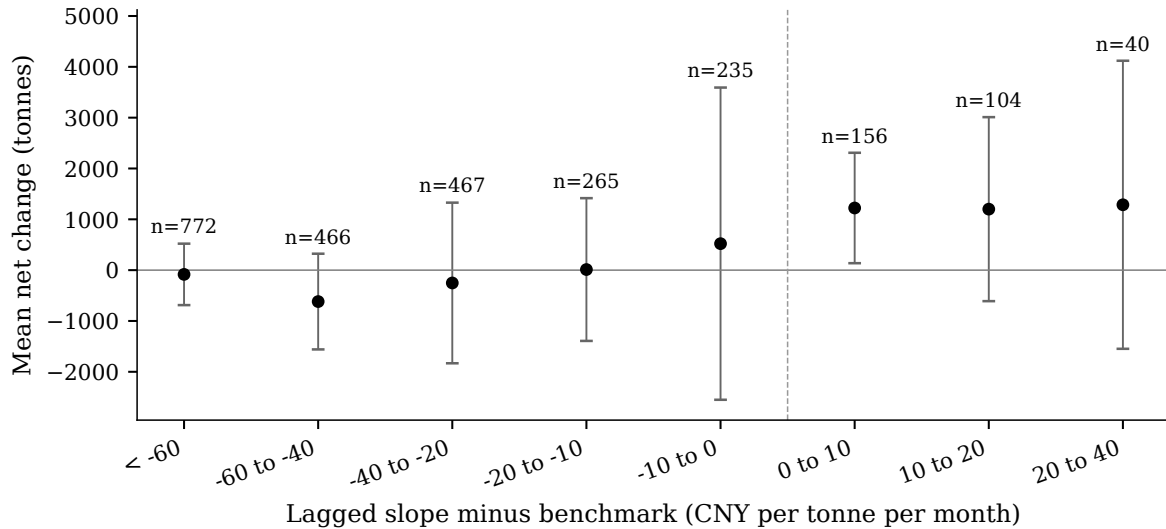


Fig. 2 Net warrant changes across bins of the lagged benchmark gap

Notes. Points are mean ΔW_t in fixed bins of g_{t-1} , measured in CNY per tonne per month. Labels give observation counts. Bars are pointwise 95 percent $t(G-1)$ approximation intervals using CRV1 year-clustered standard errors and each bin's contributing year count. They are neither simultaneous intervals nor Webb-inverted intervals. The dashed vertical line separates bins below and above zero. The bin above 40 is empty. Source: SHFE settlements and daily warrant totals.

Table 3 reports the local jump estimates. The full-sample piecewise regression gives 1,479 tonnes, with p-value 0.080. The narrowest window gives 1,696 tonnes, while the wider windows give 1,006 and 705 tonnes. Their magnitudes and precision vary across windows. The associated changes in slope are also imprecise, as reported in Online Resource 1, Section C. Taken together, these diagnostics leave the response at the precise benchmark unresolved. They do not support presenting the chosen line as an identified registration trigger.

Table 3 Local benchmark and crossing diagnostics

Window / diagnostic	Difference	SE	Webb p	N	Years
Full sample	1,479	783	0.080	2,505	11
± 10	1,696	841	0.091	391	8
± 20	1,006	1,073	0.419	760	9
± 40	705	961	0.479	1,267	11
Five-session crossing	3,737	4,420	0.418	2,501	11

Notes. The first four rows report the jump coefficient in equation (6); windows are CNY/tonne/month around the lagged zero gap. The final row compares five-session cumulative net changes beginning one session after a crossing signal with those on other dates. It uses 19 crossings. Its outcome horizon differs from the daily local tests. All differences and standard errors are tonnes. These are association diagnostics, not regression-discontinuity treatment estimates. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

The 19 crossing events produce a five-session cumulative mean of 3,770 tonnes, compared with 32 tonnes on other eligible dates. The difference is 3,737 tonnes, with standard error 4,420 and p-value 0.418. The response is therefore imprecisely estimated. Its five-session horizon differs from the one-session state comparison, and overlapping horizons share observations. The available evidence cannot determine whether imprecision is attributable mainly to the number of crossings, outcome variability, the event definition, or an absent discrete response.

An illustration from June–July 2018 reinforces the distinction between state and event interpretations. Daily warrants increase by 4,794 tonnes on 13 July and decline by 515 tonnes on 16 July, the July contract’s final trading date. Thus, an above-benchmark stretch can contain both net additions and net reductions close to expiry. Neither the

daily signs nor a cumulative change reveal whether metal physically left storage after delivery. The spell's signal dates and subsequent response dates are recorded separately in Online Resource 1, Section C.

4.4 What the findings can establish

The state comparison provides a useful empirical link between carry conditions and net warrant accumulation in the SHFE aluminum delivery system. Its limitation is equally specific: public net totals cannot tell whether an observed difference comes from more registrations, fewer cancellations, or both. The daily specification also lacks a simultaneous total-stock flow, so it cannot isolate conversion of existing unregistered metal within the same warehouses. The auxiliary weekly levels regression addresses a different question and uses the same underlying exchange sources; it is not an independent replication of the daily flow result.

The reverse regressions in Online Resource 1, Section E relate subsequent slope changes to lagged warrant changes and levels. Their estimates are imprecise under the reported inference. That outcome neither excludes feedback into futures prices nor establishes one-way causality. More generally, dated price and inventory observations can share expected demand, financing, and logistical conditions. Identifying a causal registration mechanism would require additional variation and observations of the relevant gross flows. Identifying an arbitrage boundary would additionally require a specific tradable price pair, matched financing tenor, and the full costs and constraints faced by a holder.

5 Conclusion

This study provides evidence on the association between carrying conditions and subsequent net warrant accumulation in SHFE aluminum. Above a transparent financing-and-storage benchmark, average next-session net warrant changes are higher in the unadjusted comparison. The positive sign persists in defined sensitivity specifications, while precision weakens when recent net changes, year effects, and delivery timing enter together. Local and crossing-event diagnostics leave a discrete response at the benchmark unresolved.

The contribution is to distinguish a broad state relationship in a metal delivery system from claims about a particular registration transaction. Outstanding warrants are an observable delivery-eligible stock, but their net change combines gross additions and cancellations. The fitted curve and cost proxy likewise describe carrying conditions without establishing an executable arbitrage return. Gross-flow records, matched transaction prices, and a fuller historical cost history would make those mechanisms more directly testable. With the available data, the defensible interpretation is a conditional inventory-adjustment association, whose economic content and empirical limits should be evaluated together.

Statements and Declarations

Funding

No funding was received for conducting this study.

Competing interests

The author has no relevant financial or non-financial interests to disclose.

Ethics approval and consent

This study uses aggregate exchange reports and interest-rate series. It does not involve human participants, personal data, or animals. Ethics approval and consent to participate are not applicable.

Data availability

The data supporting the reported calculations are supplied in Online Resource 2, together with source and access information. The materials include the cached SHFE contract and daily warrant series, 496 original weekly reports, selected original daily source records, and the redistributed SHIBOR series. Original exchange daily files are only partially archived, and historical SHIBOR observations have not been independently reconciled with the official benchmark archive. Historical warehouse-specific fees and separate gross registrations and cancellations are not available. Access and reuse of third-party source data remain subject to the source providers' terms.

Code availability

The analysis code, configuration, dependency information, and instructions for rebuilding the reported numerical results, tables, and figures from the bundled inputs are supplied in Online Resource 2. Rebuilding these results requires no live data query.

Supplementary information

Online Resource 1 contains input reconciliation, historical calendar details, sensitivity specifications, local and crossing-event diagnostics, auxiliary weekly evidence, and reverse predictive regressions. Online Resource 2 contains the supporting data, code, and reproducibility documentation.

References

- Brennan MJ (1958) The supply of storage. *American Economic Review* 48(1):50–72. URL <https://www.jstor.org/stable/1812340>
- Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90(3):414–427. <https://doi.org/10.1162/rest.90.3.414>
- Eastmoney (2026) Shanghai interbank offered rate: One-month renminbi series. Redistributed daily SHIBOR quotes; one-month CNY, annual percent, URL <https://data.eastmoney.com/shibor/shibor/001,CNY,201.html>, source interface accessed 1 October 2026
- Fama EF, French KR (1987) Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *The Journal of Business* 60(1):55–73. <https://doi.org/10.1086/296385>
- Fernandes VMO, Kunda EL, Robe MA (2025) Commodity futures deliveries: Theory and evidence from the US corn market. *Journal of Futures Markets* 45(7):844–876. <https://doi.org/10.1002/fut.22585>

- Garcia P, Irwin SH, Smith A (2015) Futures market failure? *American Journal of Agricultural Economics* 97(1):40–64. <https://doi.org/10.1093/ajae/aau067>
- Geman H, Scheiber M (2017) Recent experiences of copper on the Shanghai futures exchange: Some lessons for warehouse monitoring. *Resources Policy* 54:130–136. <https://doi.org/10.1016/j.resourpol.2017.09.010>
- Geman H, Smith WO (2013) Theory of storage, inventory and volatility in the LME base metals. *Resources Policy* 38(1):18–28. <https://doi.org/10.1016/j.resourpol.2012.06.014>
- Joseph K, Irwin SH, Garcia P (2016) Commodity storage under backwardation: Does the Working curve still work? *Applied Economic Perspectives and Policy* 38(1):152–173. <https://doi.org/10.1093/aep/ppv011>
- Kim S, Kim J, Heo E (2017) Convenience yield of accessible inventories and imports: A case study of the Chinese copper market. *Resources Policy* 52:277–283. <https://doi.org/10.1016/j.resourpol.2017.04.003>
- MacKinnon JG, Webb MD (2017) Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics* 32(2):233–254. <https://doi.org/10.1002/jae.2508>
- MacKinnon JG, Webb MD (2018) The wild bootstrap for few (treated) clusters. *The Econometrics Journal* 21(2):114–135. <https://doi.org/10.1111/ectj.12107>
- Shanghai Futures Exchange (2019) Non-ferrous metals futures contract trading manual, 2019 edition. Official Chinese manual, aluminum contract and delivery provisions, printed pp. 27 and 36; warehouse charges, p. 68, URL <https://www.shfe.com.cn/upload/20190919/1568880174521.pdf>, accessed 1 October 2026
- Shanghai Futures Exchange (2025) Daily futures market reports, daily warrant reports, and weekly warehouse stock reports. Historical aluminum series, 2015–2025, URL <https://www.shfe.com.cn/data/tradedata/future/dailydata/>, contract and daily warrant series retrieved through AKShare; weekly source reports archived locally; selected original daily records verified 1 October 2026

353 Shanghai Futures Exchange (2026) Aluminum futures delivery warehouses and
 354 charges. Official warehouse list and fee table, September 2026 snapshot, URL
 355 [https://www.shfe.com.cn/products/futures/metal/nonferrousmetal/al_f/warehouses/](https://www.shfe.com.cn/products/futures/metal/nonferrousmetal/al_f/warehouses/202609/P020260911391299421888.pdf)
 356 [202609/P020260911391299421888.pdf](https://www.shfe.com.cn/products/futures/metal/nonferrousmetal/al_f/warehouses/202609/P020260911391299421888.pdf), accessed 1 October 2026; snapshot version does
 357 not establish applicability throughout 2015–2025

358 Uría R, Williams J (2007) The supply of storage for natural gas in California. *The Energy*
 359 *Journal* 28(3):31–50. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol28-No3-3>

360 Webb MD (2023) Reworking wild bootstrap-based inference for clustered errors. *Canadian*
 361 *Journal of Economics* 56(3):839–858. <https://doi.org/10.1111/caje.12661>

362 Working H (1949) The theory of price of storage. *American Economic Review* 39(6):1254–
 363 1262. URL <https://www.jstor.org/stable/1816601>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1.pdf](#)