Internet Appendix for "Anomalies on a Schedule: Month-End Liquidity Demand and the Factor Zoo"

Daniel Nathan* Matti Suominen†

This version: August 2026

Abstract

This Internet Appendix reports supplementary evidence for "Anomalies on a Schedule: Month-End Liquidity Demand and the Factor Zoo." The evidence is organized around the paper's main identifying claims. First, we show that the PreTOM concentration is not an artifact of window choice, sample period, or return adjustment. Second, we test alternative dispensability measures and show that distance below the 52-week high remains the best single-characteristic proxy. Third, we report additional T+1 settlement-reform, ETF, stock-level, and within-leg evidence. Finally, we address factor redundancy, external-library replication, multiple testing, and full factor-level classifications.

Roadmap.

*Hong Kong Polytechnic University. Email: daniel.nathan@polyu.edu.hk
†Aalto University School of Business. Email: matti.suominen@aalto.fi

Question	*Evidence*	*Result*
Window choice	Permutation and bootstrap window tests	The cash-demand window is where factor returns disperse most
Dispensability proxy	Alternative constructions and multiple-testing bootstrap	Distance below the 52-week high remains the best single-characteristic proxy
Settlement timing	Inference comparison, placebo dates, pre-period starts	Sign and timing unchanged
Selling pressure	TAQ order flow, stock-level, and within-leg evidence	Dispensable stocks face extra seller-initiated trading before month-end
Factor construction	Official JKP portfolios and independent libraries	The asymmetry holds under official JKP and external libraries
Factor redundancy	PCA, pruning, theme weighting, multiple testing	Redundancy is symmetric across windows; the return concentration is not

Table IA.1: Variable Definitions

Object	Definition	First use
$\delta_{i,m}$	Stock i's dispensability score at month-end m: the within-month z-score of distance below the 52-week high (higher = more dispensable)	Section 3.1
d_f	Factor f's average short-minus-long exposure to stock-level dispensability (positive: the factor profits when dispensable stocks are sold)	Section 3.1
LDS_m	Monthly return spread measuring the realized PreTOM price concession on dispensable stocks	Section 4.1
β_f^{LDS}	Time-series exposure of factor f to LDS_m	Section 4.2
d_e	ETF e's holdings-weighted exposure to stock-level dispensability	Section 5.1
PreTOM	Trading days $\tau-9$ through $\tau-4$, where τ is the last trading day of the month	Section 2.2
Rest	Trading days outside PreTOM	Section 2.2
Post	Trading days $\tau-3$ through $\tau+3$	Section 2.2

IA.1 Bootstrap and Permutation Inference

For exact reproducibility, every bootstrap and permutation test fixes the random seed. In the Python pipeline the seeds are `20260513` for the cross-factor d_f bootstraps and the robustness battery (correlation pruning, effective rank, idiosyncratic- variance weighting, theme weighting, multiple testing, external catalogs), `20260613` for the characteristic-ranking bootstrap, `20260707` for the random-window placebo (Table IA.2), `20260507` for the M2-variant and Stambaugh–Yuan permutation tests, `20260517`, `20260518`, and `20260523` for the reversal and ρ_f tests, and `42` for the factor-leg TAQ test; the Stata replication and the orchestrated horse-race, external- catalog, and T+1 bootstraps use `set seed 20260507`. At $B = 5{,}000$ month-block draws the resulting intervals and t-statistics are stable across seeds to two decimal places.

Table IA.2: Random-Window Placebo Test of PreTOM Concentration

Test statistic	Actual	Null mean (sd)	Null 95th	p-value
Mean absolute PreTOM share	0.445	0.339 (0.035)	0.401	0.002
Cross-factor variance ratio	3.88	1.33 (0.59)	2.47	0.003

Notes: 10,000 placebo draws (seed `20260707`). In each draw we select 6 random trading days per month as a placebo window and recompute both statistics. Per-factor window means average within-month window means across months. The variance ratio is $\mathrm{Var}_f(\bar{R}_f^{\mathrm{Pre}})/\mathrm{Var}_f(\bar{R}_f^{\mathrm{Rest}})$ across factors; the mean absolute share is the cross-factor average of $|6\,\bar{R}_f^{\mathrm{Pre}}|/(|6\,\bar{R}_f^{\mathrm{Pre}}| + |15\,\bar{R}_f^{\mathrm{Rest}}|)$. Raw VW long–short decile returns, NYSE breakpoints, 1963–2025. p-values are one-sided shares of placebo draws at or above the actual statistic.

The anchor cross-sectional specification uses inverse-variance WLS, where the weight on each factor is the inverse of the analytical variance of Δ_f. Table IA.4 reports the slope under alternative weighting choices: equal-weighted OLS, WLS with weights winsorized at 1/99 and 5/95, WLS with weight caps at 10× and 20× the median weight, a theme-equal regression, and the range of leave-one-theme-out and leave-one-factor-out WLS slopes. The estimates remain positive throughout, so the settlement-shift sign does not depend on the weighting scheme.

The estimate is likewise insensitive to the start of the pre-reform window used in the bootstrap. Re-running the bootstrap with pre-reform starts of 1980, 2000, 2010, and 2017-09 (the T+2 settlement era) leaves the anchor $\hat{b}$ varying by less than 10% (range +29.0 to +32.0) and bootstrap t by less than 0.13 (range +1.72 to +1.84). Power is bound by the fixed 19-month post-reform sample, not by pre-reform length.

Table IA.3: T+1 Cross-Sectional Test: Inference Comparison

Spec	Slope (bps)	Inference	Statistic	One-sided p
OLS $\hat{\Delta}_f \sim d_f$	+37.2	Theme-clustered analytical	$t = +2.01$	0.022
	+37.2	Month-block bootstrap	$t = +1.44$	0.075
	+37.2	Placebo reform-date randomization	95.9th pctile	0.042
WLS $\hat{\Delta}_f \sim d_f$	+52.2	Theme-clustered analytical	$t = +2.95$	0.002
	+52.2	Month-block bootstrap	$t = +1.65$	0.050
	+52.2	Placebo reform-date randomization	**99.1th pctile**	**0.010**

Notes: Three inference frameworks for the same cross-sectional test: signed factor return DiD $\hat{\Delta}_f = \hat{\pi}_{f,\tau-3} - \hat{\pi}_{f,\tau-4}$ regressed on dispensability exposure d_f across 152 JKP factors. *Theme-clustered analytical* groups factors by JKP theme (14 clusters) and applies the standard cluster-robust SE; it captures within-theme correlation but conditions on the per-factor point estimates. *Month-block bootstrap* resamples the pre- and post-reform months separately (B = 5,000), recomputing $\hat{\Delta}_f$ inside each draw; it additionally captures first-stage estimation uncertainty from the 19-month post-reform window, with $t = \hat{b}/\mathrm{SE}_{\mathrm{boot}}$. *Placebo reform-date randomization* draws 1,000 pseudo-reform dates uniformly from 1996–2015, so every placebo post-window lies inside the T+3 era and all windows end by April 2017, well clear of both the September 2017 T+3→T+2 and May 2024 T+2→T+1 transitions (pre-windows of the earliest draws reach back to 1989 and touch the pre-June-1995 T+5 era); the same DiD machinery is applied to each placebo and the actual May 2024 slope is compared to the placebo distribution. All three rows within each panel report the same anchor slope (OLS +37.2, WLS +52.2); the inference methods differ only in how they estimate its standard error. The three procedures address different sources of uncertainty; the placebo distribution preserves the cross-sectional dependence structure of the data, and we use it as the inference for the factor-level T+1 test.

Table IA.4: T+1 Cross-Sectional Test: Weighting Robustness

Specification	N	Slope (bps)	t	Weight rule
OLS (equal-weighted)	152	+37.23	+3.40	1 / factor
WLS, cell-mean dispersion SE	152	+56.24	+4.61	$1/\widehat{\mathrm{SE}}(\Delta_f)^2$ (from cell-mean SD)
WLS, weights winsorized at 1/99	152	+56.20	+4.60	clip(w, p1, p99)
WLS, weights winsorized at 5/95	152	+55.41	+4.52	clip(w, p5, p95)
WLS, weights capped at 10x median	152	+56.24	+4.61	$\min(w, 10\,\mathrm{median}(w))$
WLS, weights capped at 20x median	152	+56.24	+4.61	$\min(w, 20\,\mathrm{median}(w))$
Theme-equal regression	14	+72.43	+2.34	theme means, equal-weighted
Leave-one-theme-out WLS, min	152	+47.11	—	min over 14 themes
Leave-one-theme-out WLS, max	152	+71.64	—	max over 14 themes
Leave-one-factor-out WLS, min	152	+53.45	—	min over 152 factors
Leave-one-factor-out WLS, max	152	+62.17	—	max over 152 factors

Notes: Cross-sectional regression of per-factor signed DiD $\Delta_f = \pi_{f,\tau-3} - \pi_{f,\tau-4}$ on the dispensability exposure d_f across 152 JKP factors (the 52-week-high self-loading factor excluded). The WLS weight in every row of this table is $1/\widehat{\mathrm{SE}}(\Delta_f)^2$, where $\widehat{\mathrm{SE}}(\Delta_f)$ is computed from the dispersion of pre/post-period cell means for $\tau-4$ and $\tau-3$. This is a different variance estimator from the main-paper Table 8 Panel A WLS anchor (+52.2, $t = +2.95$), which uses the analytical inverse variance of Δ_f implied by the per-factor day-level DiD regression. The two estimators yield slopes in the same direction and same order of magnitude (+52.2 vs. +56.24) but are not identical because they place different weights on factors with sparse post-reform months. "Theme-equal" first averages Δ_f and d_f within each JKP theme (14 themes) and runs the regression on theme means. Leave-one-theme-out and leave-one-factor-out report the range of WLS slopes across all single-element exclusions. Pre-period: T+2 era (Sept 5, 2017–Apr 30, 2024). Post-period: T+1 era (June 1, 2024–Dec 31, 2025). May 2024 dropped.

IA.2 Alternative Dispensability Constructions

This section asks whether the main result depends on the exact dispensability proxy. The main paper uses one characteristic: distance below the 52-week high. Composite measures can raise in-sample fit, but they also risk importing profitability, quality, or earnings-streak premia that exist outside PreTOM. We therefore treat the 52-week-high measure as the baseline and report composites as robustness checks. Table IA.5 reports the cross-factor R^2 and clustered t-statistic from the main-paper specification across alternative dispensability constructions.

Table IA.5: Alternative Dispensability Measures

	PreTOM		Rest	
Construction	R^2	t	R^2	t
Panel A: Matched robustness baseline and components				
Price / 52-week high only — matched robustness baseline	**0.586**	**8.02**	**0.025**	**1.04**
Financial safety (QMJ-safety) alone	0.473	5.11	0.000	−0.15
Panel B: Composite measures (robustness only)				
52-week high + QMJ-safety	0.667	9.40	0.005	0.59
52-week high + Z-score	0.562	7.19	0.017	0.82
52-week high + F-score	0.591	11.39	0.079	2.01
52-week high + QMJ	0.656	11.04	0.041	1.64
52-week high + QMJ-safety + profitability	0.707	9.21	0.030	1.39

Notes: The main paper uses the single-characteristic 52-week-high specification. Composite measures are reported only as robustness checks. We do not use them as the baseline because they can import non-PreTOM profitability or quality premia into the dispensability measure. Each row reports a cross-factor regression of average daily PreTOM (Rest) long-short return on the indicated dispensability construction. All 153 JKP factors. t-statistics clustered by JKP theme (14 clusters). NYSE breakpoints, VW deciles, market-adjusted, 1980–Jan 2025.

B.1. Stock-Level Residualization of the Dispensability Score

Table IA.6 asks whether the 52-week-high measure is only a repackaging of standard characteristics. Each month, we residualize the stock-level dispensability measure on reversal, prior 11-month return, market beta, idiosyncratic volatility, and QMJ-safety, then rebuild the factor-level short-minus-long loading from the residual. The residualized loading continues to explain a large share of the PreTOM cross-section: the PreTOM R^2 is 0.386 with $t = +6.75$, compared with 0.724 for the baseline measure. About 53% of the baseline PreTOM explanatory power remains after residualization. The Rest-window fit remains small

and statistically insignificant, with $R^2 = 0.009$ and $t = +0.66$. Standard characteristics attenuate the 52-week-high signal but do not explain the relation between dispensability exposure and PreTOM factor returns. We retain the single-characteristic specification as the baseline measure for its parsimony and direct economic interpretation; the residualized version is reported here as a robustness check.

Table IA.6: Residualized d_f: Standard-Characteristic Controls

	PreTOM		Rest	
Dispensability measure	R^2	t	R^2	t
Baseline d_f (52-week high only)	0.724	+12.41	0.015	+0.82
Residualized d_f	**0.386**	**+6.75**	0.009	+0.66

Notes: The baseline d_f uses the stock-level distance-from-52-week-high measure. The residualized version first projects the stock-level measure each month on reversal, prior 11-month return, market beta, idiosyncratic volatility, and QMJ-safety, and then constructs d_f from the residual. The table reports cross-factor regressions of mean factor returns on each loading separately in PreTOM and Rest. Factor-level loadings use value-weighted NYSE-breakpoint D1–D10 averages of the stock-level score; the single-characteristic 52-week-high factor is excluded (self-loading). t-statistics clustered by JKP theme (14 clusters). Sample: 1963–Dec 2025.

Table IA.7 extends this exercise to the time-series-momentum sell signal, an indicator for a negative trailing twelve-month return. Dispensable stocks overlap heavily with the signal, yet the exposure is not subsumed by it. Residualizing on the sell indicator alone leaves the PreTOM fit at 60%, and a significant component (29%, $t = +5.3$) survives adding the indicator to the full battery.

Table IA.7: Time-Series Momentum and the Dispensability Exposure

	PreTOM		Rest	
Characteristics residualized out	R^2	t	R^2	t
TSM sell indicator only	60%	+11.8	0%	−0.3
Full battery + TSM sell indicator	29%	+5.3	0%	+0.3

Notes: Each row residualizes the stock-level distance-from-52-week-high score on the listed characteristics by monthly cross-sectional regression, then rebuilds the factor-level exposure as the value-weighted NYSE-breakpoint D1-minus-D10 average of the residualized score and regresses each factor's intended-signed average daily long–short return on it, across the 152 factors (the 52-week-high factor excluded), separately in PreTOM and Rest. The construction matches Table IA.6. The TSM sell indicator equals one for a negative trailing twelve-month return. The battery is short-term reversal, prior 11-month return, market beta, idiosyncratic volatility, and QMJ-safety. t-statistics are theme-clustered. 1963–2025.

IA.3 The Liquidity-Demand Shock

This section collects the exhibits describing the monthly liquidity-demand shock LDS_m of equation (11): its magnitude across intra-month windows, its robustness to rebuilding the shock from stocks the test factor does not hold, whether its size can be forecast in advance, and how its price impact depends on the state of the market.

Window means and holdings exclusion

Table IA.8 reports the mean spread between the least- and most-dispensable deciles in each intra-month window. Table IA.9 reports the by-theme correlations between the full-universe shock and the shocks rebuilt after excluding the extreme-leg holdings of every factor in the test factor's theme, together with alternative exclusion rules and per-month exposure specifications.

Table IA.8: Unconditional Mean of LDS_m by Trading-Day Window

	Mean (bps/day)	t-stat (HC1)	N days
Mid-month placebo $[\tau{-}15, \tau{-}10]$	$+1.67$	$(+0.72)$	4,529
PreTOM $[\tau{-}9, \tau{-}4]$	$+8.93^{***}$	$(+4.14)$	4,530
Post $[\tau{-}3, \tau{+}3]$	-4.61^{**}	(-2.23)	5,285

Notes: Unconditional mean of LDS_m, the return of the least-dispensable stock decile minus the most-dispensable (equation 11), in bps/day across three within-month windows, 1963–2025. HC1 heteroskedasticity-robust standard errors. *,** ,*** denote 10%, 5%, 1% significance.

Forecastability of the shock's size

Table IA.10 asks whether the size of a given month's shock can be forecast from information available beforehand, and whether realized fund flows explain it after the fact. It backs the timing-versus-magnitude discussion of Section 4.1.

State dependence with liquidity and flow controls

Table IA.11 is the appendix companion to main-text Table 7. Panel A adds liquidity and flow controls to the net-seller-pressure interaction with predetermined volatility; Panel B replaces the state variable with mutual-fund outflows.

Table IA.9: Alternative LDS Exclusion Rules

Panel A: Exclusion summary and panel-regression coefficients

	Correlation with full LDS			PreTOM		Rest		
Exclusion rule	mean	min	max	β	t	β	t	N
Full LDS (no exclusion)	1.00	1.00	1.00	+ 0.505	+16.71	+ 0.016	+0.66	113,911
Theme-wide own-stock exclusion	0.78	0.43	0.98	+ 0.424	+11.86	+ 0.035	+1.31	113,911

Panel B: Theme-wide exclusion – correlation with full LDS *by JKP theme*

Theme	Corr.	Theme	Corr.	Theme	Corr.	Theme	Corr.
Other	0.43	Profitability	0.73	Profit Growth	0.84	ST Reversal	0.94
Low Risk	0.60	Momentum	0.79	Quality	0.85	LT Reversal	0.98
Value	0.61	Debt Issuance	0.81	Accruals	0.87		
Investment	0.72	Seasonality	0.84	Size	0.93		

Notes: Panel A reports the cross-factor panel coefficient on $d_f \times$ LDS under two exclusion rules. "Full LDS" uses the full CRSP universe; "Theme-wide own-stock exclusion" removes the union of D_1 and D_{10} holdings across all factors in the same JKP theme as f each month. After each exclusion, the remaining universe is re-ranked by dispensability and the NYSE-breakpoint deciles are recomputed before constructing the $D_{10} - D_1$ PreTOM spread. "Correlation with full LDS" in Panel A averages the per-theme correlation across the matched 1963–2025 sample. Panel B reports the per-theme correlation between the theme-excluded LDS series and the full-universe LDS. All specifications: factor and month fixed effects, standard errors two-way clustered by factor and month, matched sample 1963–2025. The 52-week-high characteristic (which defines d_f) is excluded from the panel throughout.

Table IA.10: The Size of the Month-End Liquidity Event Is Not Forecastable

Panel A: Out-of-sample forecasts of the event's size				
Dependent variable	Predictors (frozen ex ante)	Sample	N_{OOS}	OOS R^2
$\text{LDS}_m^{\text{PreTOM}}$	Dividends due, calendar dummies, settlement lag	1963–2025	623	−0.02
$\text{LDS}_m^{\text{PreTOM}}$	+ lagged MF flows, market return, LDS_{m-1}	1990–2025	162	−0.13
$\text{LDS}_m^{\text{PreTOM}}$	Lagged aggregate MF flow, $m-1$	2000–2023	221	−0.07
$\text{NetSell}_m^{\text{PreTOM}}$	Dividends due, calendar dummies, settlement lag	2003–2022	110	+0.39
$\text{NetSell}_m^{\text{PreTOM}}$	+ lagged MF flows, market return, LDS_{m-1}	2003–2022	110	+0.36
$\text{Flow}_m^{\text{PreTOM}}$	Lagged aggregate MF flow, $m-1$	2000–2023	221	−0.00
Panel B: Realized PreTOM fund flows and same-month outcomes				
Regression		Slope	t	R^2
LDS_m on realized $\text{Flow}_m^{\text{PreTOM}}$		−4.31	−1.84	0.010
$\text{NetSell}_m^{\text{PreTOM}}$ on realized $\text{Flow}_m^{\text{PreTOM}}$		−0.025	−0.55	0.006

Notes: Panel A reports Campbell–Thompson out-of-sample R^2 against the expanding historical mean. Each forecast for month m uses only data dated $m-1$ or earlier, and is estimated on all earlier months, starting once 120 months of history are available for the calendar specifications and 60 for the shorter flow samples. All three dependent variables are measured over the six PreTOM days of month m: $\text{LDS}_m^{\text{PreTOM}}$ is the dispensability spread of equation (11), $\text{NetSell}_m^{\text{PreTOM}}$ is net seller-initiated dollar volume (WRDS Intraday Indicators, share of classified volume), and $\text{Flow}_m^{\text{PreTOM}}$ is the total net flow into U.S. equity mutual funds, scaled by market capitalization and aggregated from fund-level Morningstar data, where positive values are net inflows. The predictors are month-level variables known at the close of $m-1$, not return windows. The first set describes when cash obligations fall due in month m: dividends payable between $\tau-9$ and month-end that were already declared by the end of $m-1$, quarter-end, year-end, and April dummies, and the settlement lag in force. The augmented set adds the same aggregate MF flow series lagged to $m-1$, the $m-1$ market return, and LDS_{m-1}. Both were fixed before estimation. The positive OOS R^2 for $\text{NetSell}_m^{\text{PreTOM}}$ reflects a single regime shift at the 2017 T+3-to-T+2 settlement change that the expanding-mean benchmark cannot track: within settlement regimes the OOS R^2 is −0.04 (T+3 era) and +0.82 (T+2 era), and no predictor other than the settlement lag is significant. Panel B regresses each outcome on the realized (ex-post) window flow, with Newey–West(12) standard errors. Because window flow is itself unforecastable (Panel A), the realized flow is almost entirely unexpected: stripping out a real-time forecast built from lagged flows and quarter-end dummies leaves a shock whose standard deviation is 0.99 times that of the raw series, and regressing $\text{LDS}_m^{\text{PreTOM}}$ on that shock gives −5.2 ($t = -2.08$). The distinction between realized and unexpected flow is therefore immaterial here.

Table IA.11: State Dependence: Liquidity Controls and Flow Amplification

Dependent variable: $\text{LDS}_m^{\text{PreTOM}}$ (bps/day)					
Panel A: Liquidity and flow controls					
	(1)	(2)	(3)	(4)	(5)
Net seller pressure, NSP_m	+19.4	+19.4	+19.3	+19.7	+18.3
	(+2.30)	(+2.29)	(+2.26)	(+2.22)	(+2.30)
VIX_{m-1}	−11.6	−11.5	−11.2	−12.1	−3.7
	(−1.11)	(−1.11)	(−1.01)	(−1.16)	(−0.42)
$\text{NSP}_m \times \text{VIX}_{m-1}$	+36.5	+36.5	+36.7	+36.1	+32.6
	(+3.18)	(+3.21)	(+3.24)	(+3.17)	(+3.39)
Pástor–Stambaugh liquidity	No	Yes	No	No	Yes
Amihud illiquidity	No	No	Yes	No	Yes
Mutual-fund outflows	No	No	No	Yes	Yes
Market return	No	No	No	No	Yes
N (months)	230	230	230	230	230
R^2	0.203	0.203	0.203	0.204	0.286
Panel B: Amplification by mutual-fund outflows					
	Window outflows		Prior-month	Rest	
	(1)	(2)	(3)	(4)	
$\text{NSP}_m \times$ MF outflows	+27.1	+23.4	+14.1	−8.3	
	(+2.70)	(+2.98)	(+1.55)	(−1.66)	
$\text{NSP}_m \times \text{VIX}_{m-1}$	—	+31.0	—	—	
		(+3.77)			
N (months)	230	230	230	230	
R^2	0.150	0.257	0.076	0.044	

Notes. Companion to Table 7, adding liquidity and flow controls one at a time and then jointly. The dependent variable is the monthly PreTOM concession $\text{LDS}_m^{\text{PreTOM}}$ of equation (11). NSP_m is the market-wide net-seller-pressure spread of equation (13) and VIX_{m-1} is the VIX level at the prior month-end, so the state is fixed before the window opens. Controls are the Pástor–Stambaugh traded liquidity innovation, log Amihud illiquidity, aggregate mutual-fund outflows over the PreTOM days, and the contemporaneous market return; all regressors are standardized, so coefficients are basis points per day per standard deviation. The interaction of interest, $\text{NSP}_m \times \text{VIX}_{m-1}$, is stable across the ladder and retains significance with all four controls included (column 5). Panel B replaces the state variable with aggregate mutual-fund outflows over the same PreTOM days, standardized, and asks whether a given amount of net seller pressure moves prices more when redemptions are heavy. Column 1 enters the flow interaction alone, column 2 enters it jointly with the volatility interaction, column 3 replaces window outflows with the prior month's window outflows so that the conditioner is predetermined, and column 4 is the Rest-window placebo. Sample: 2003–2022, the TAQ era. Newey–West(5) standard errors in parentheses.

IA.4 The PreTOM Window: Structure and Alternatives

This section describes the window itself: how tightly the cross-section coheres across its days, how broadly the concentration holds across subperiods, and how the main result responds to alternative window definitions.

Cross-day coherence and subperiod breadth

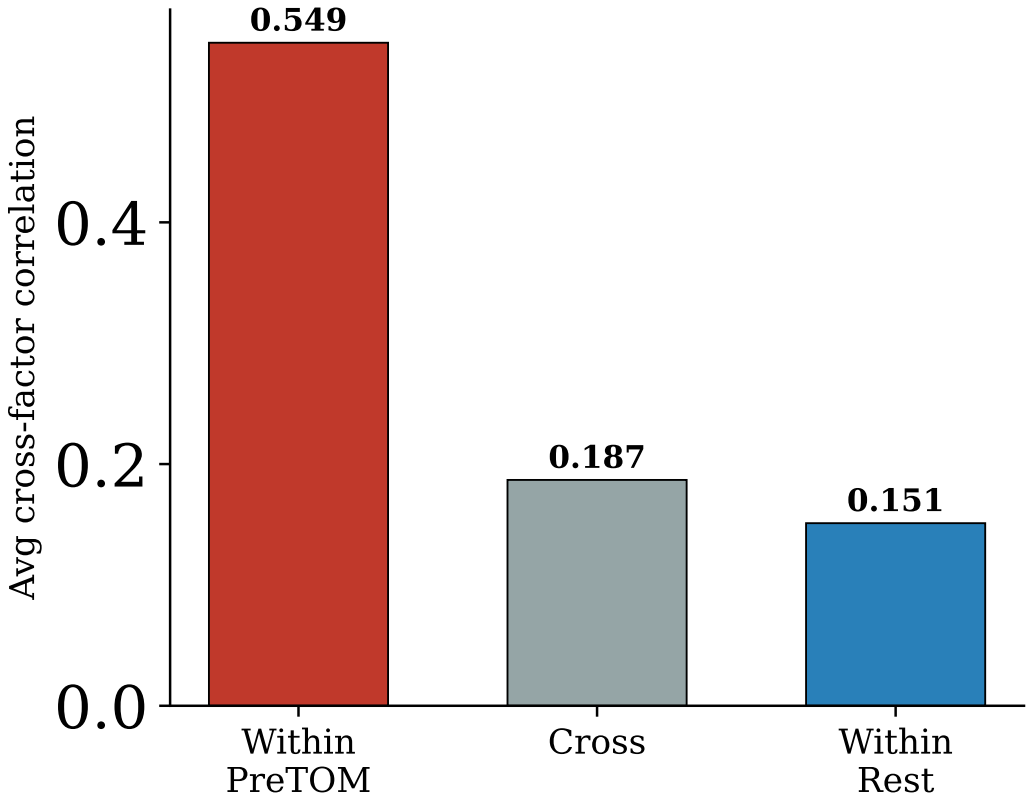

Figure IA.1: Cross-factor pairwise correlation between trading days, 1963–Dec 2025. For each trading-day position from month-end ($\tau-0$ through $\tau-20$) we average daily long–short returns across all calendar months, giving a 153-dimensional cross-factor vector for each position; we then correlate these vectors across position pairs and average within categories. Bars show the average within-PreTOM correlation (pairs of positions both inside the $[\tau-9, \tau-4]$ window), the average within-Rest correlation (pairs of positions both outside), and the average cross correlation (one position inside and one outside).

Table IA.12: Subperiod Stability of |Pre| > |Rest| Count

Sample	N	|Pre| > |Rest|	Share
Full (1963–2025)	153	113	73.9%
1963–1993	153	112	73.2%
1994–2025	153	107	69.9%

Notes: For each of the 153 JKP factors and each subperiod, we compute the average daily long–short return inside the PreTOM window ($\tau-9$ to $\tau-4$) and outside it. The third column counts the number of factors whose absolute PreTOM daily return exceeds their absolute Rest daily return. The count is similar in both halves of the sample. Because a six-day window's mean return is noisier than its roughly fifteen-day complement, this count exceeds 153×0.5 even when no day is special; we therefore report it as a descriptive breadth statistic, and the formal random-window placebo tests of the window are in Table IA.2. VW deciles, NYSE breakpoints.

T+1 Settlement Reform: Permutation and Weighting Inference

The T+1 cross-sectional test of Section 5 regresses the per-factor DiD on the raw dispensability gap $\widetilde{d}_f$, using raw $D10 - D1$ returns. This subsection reports four checks on its inference: a placebo permutation test of the reform date, the analytical-versus-bootstrap inference comparison, the sensitivity of the cross-sectional slope to the factor-weighting scheme, and the robustness of the bootstrap to the pre-reform window start. Table IA.13 reports the per-factor timing shift averaged within JKP themes, alongside the theme-average dispensability gap. These means are descriptive. The identifying test is the cross-factor slope in Table 8, which is estimated with an intercept that absorbs the reform's common timing-shift component, so a theme's raw mean shift need not share the sign of its mean gap.

E.1. Alternative Window Definitions

We test the sensitivity of our results to the PreTOM window definition. Table IA.14 reports the d_f cross-sectional R^2 for windows of 5, 6, and 7 days.

E.2. Pre-2002 Rest-Window Leakage Diagnostic

The within-half replication (Table IA.17) shows a small but significant Rest-window R^2 of 0.121 ($t = 2.60$) in 1980–2002 that vanishes post-2002 ($R^2 = 0.075$, $t = -1.87$). We decompose the Rest window into near-Rest ($\tau-3$ through $\tau-1$, the three days immediately following the PreTOM window) and far-Rest ($\tau-20$ through $\tau-10$) to localize the leakage (Table IA.15).

Table IA.13: T+1 Timing Shift by JKP Theme

Theme	N	Mean $\widetilde{d}_f$	Mean Δ_f (bps)
Investment	27	+0.00	+80.9
Profit Growth	5	+0.21	+66.6
Momentum	7	+1.15	+64.4
Seasonality	10	+0.18	+59.8
Profitability	21	+0.36	+54.5
Accruals	8	+0.01	+53.9
Low Risk	20	−0.36	+39.1
Other	27	−0.09	+20.1
Quality	7	+0.07	+18.2
Value	8	+0.11	−17.7
Debt Issuance	8	−0.12	−35.4
Size	2	−0.07	−64.5

Notes: Descriptive JKP-theme means of the per-factor timing shift Δ_f and of the raw dispensability gap $\widetilde{d}_f$, over the 152 factors of Table 8. These means are not the identifying test: the settlement prediction is the positive cross-factor slope of Δ_f on $\widetilde{d}_f$, estimated with an intercept, and the intercept absorbs the reform's common timing-shift component. A theme's raw mean Δ_f therefore need not share the sign of its mean $\widetilde{d}_f$. Themes with fewer than two factors are omitted.

Table IA.14: Cross-Sectional R^2 by Window Definition

Window	Days	PreTOM R^2	PreTOM t
$\tau-8$ to $\tau-4$ (5 days)	5	0.623	9.21
$\tau-9$ to $\tau-4$ (baseline)	6	0.659	9.63
$\tau-10$ to $\tau-4$ (7 days)	7	0.681	9.96

Notes: Cross-factor regression of average daily PreTOM long–short return on dispensability exposure, varying the window definition. The baseline 6-day window ($\tau-9$ through $\tau-4$) is pinned by the aggregate market's return path (Figure 2 in the main paper), not optimized over R^2.

Table IA.15: Rest-Window Leakage by Sub-Window and Period

	1980–2002		2003–Jan 2025	
Sub-window	R^2	t	R^2	t
Near-Rest ($\tau-3$ to $\tau-1$)	0.49	+5.5	—	−5.9
Far-Rest ($\tau-20$ to $\tau-10$)	0.07	—	—	—

Notes: The pre-2002 Rest significance concentrates entirely in the near-Rest window ($\tau-3$ through $\tau-1$), not in trading days far from the PreTOM boundary. After decimalization in 2001, the near-Rest coefficient reverses sign ($t = -5.9$), consistent with faster price discovery containing the institutional impact inside the six-day PreTOM window. The far-Rest R^2 of 0.07 in 1980–2002 is an order of magnitude below the near-Rest value, consistent with the pre-2002 leakage reflecting slow price adjustment near the window boundary rather than a broader month-long mechanism. Cells left blank are not reported.

E.3. Day-by-Day Return Profile

Table IA.16 reports the average signed long–short return across 153 JKP factors for each intramonth phase, decomposing the monthly return into its calendar components.

Table IA.16: Daily Return Profile: Signed Long-Short Spread by Trading-Day Position

Phase	Days	Avg LS (bps/day)	t
PreTOM ($\tau-9$ to $\tau-4$)	6	+2.51	+9.94
Recovery ($\tau-3$ and $\tau+1$)	2	+3.48	+10.92
Last day (τ)	1	−4.83	−6.83
Core Rest (remaining ∼12 d)	12	+1.46	+12.95
Monthly contributions (bps/month)			
PreTOM zone (8 days)	8	+22.0 (≈ 60% of monthly total)	
Core Rest + last day (13 days)	13	+12.7 (≈ 40% of monthly total)	

Notes: Average daily signed long-short return across 153 JKP factors, grouped by intramonth phase. Each factor is signed by its full-sample mean LS direction. "Recovery" includes $\tau-3$ (the first Rest day after PreTOM ends within the same month) and $\tau+1$ (the first trading day of the following month). Core Rest excludes $\tau-3$, $\tau+1$, and τ. Cross-factor t-statistics. VW deciles, NYSE breakpoints, 1980–Jan 2025.

IA.5 Within-Half Replication

Table IA.17 replicates the core d_f cross-sectional regression independently in each half of the sample (1980–2002 and 2003–Jan 2025), re-estimating both dispensability exposures and factor premia within each half.

Table IA.17: Within-Half Replication: $d_f \rightarrow$ Daily LS Returns

		λ_f (Pre−Rest)		PreTOM		Rest	
Period	N	R^2	t	R^2	t	R^2	t
1980–2002 (H1)	152	0.557	7.16	0.597	6.72	0.121	2.60
2003–Jan 2025 (H2)	153	0.423	9.93	0.381	6.17	0.075	−1.87

Notes: Cross-factor regressions of daily long–short return on d_f, estimated independently in each half. Dispensability exposures are recomputed within each half using only that half's stock-level characteristics and VW decile assignments; λ_f = PreTOM mean − Rest mean. t-statistics are clustered by JKP theme (14 clusters).

IA.6 Dispensability Out-of-Sample Test

Table IA.18 tests whether dispensability exposures estimated in one half of the sample predict PreTOM concentration in the other half.

Table IA.18: Dispensability Out-of-Sample Test

d_f formed in	λ_f from	R^2	t	Type
Panel A: Predicting λ_f				
1980–2002	1980–2002	0.613	—	In-sample
2003–Jan 2025	2003–Jan 2025	0.447	—	In-sample
1980–2002	2003–Jan 2025	**0.421**	**+10.47**	**Out-of-sample**
2003–Jan 2025	1980–2002	**0.575**	**+14.29**	**Out-of-sample**
Panel B: Predicting PreTOM mean				
1980–2002	2003–Jan 2025	0.508	+12.49	Out-of-sample
2003–Jan 2025	1980–2002	0.682	+18.00	Out-of-sample
Panel C: Predicting Rest mean (falsification)				
1980–2002	2003–Jan 2025	0.007	+1.04	Out-of-sample
2003–Jan 2025	1980–2002	0.012	+1.35	Out-of-sample
Panel D: d_f stability				
Corr(d_f^{H1}, d_f^{H2}) across 153factors				+0.958

Notes: d_f is formed independently in each half-sample using only that half's stock-level characteristics and VW decile assignments. λ_f is estimated independently in each half from time-series regressions on daily LS returns. t-statistics are clustered by JKP theme (14 clusters).

Interpretation. What is being held fixed in the out-of-sample test is each factor's *characteristic tilt*, d_f, not the identity of the stocks inside its deciles. The specific stocks sitting in D1 or D10 of any factor turn over continuously: momentum's worst losers in 1985 (US Steel, Bethlehem Steel) are not its worst losers in 2015 (Chesapeake Energy, J.C. Penney). What persists is the mechanical relationship between each factor's *sorting rule* and the dispensability characteristic. Sorting on past returns pulls stocks that are far from their 52-week high and financially fragile into D1 in every decade, because large recent losers are almost by definition far from their highs and frequently in distress. Sorting on book-to-market pulls similar stocks into D10. Sorting on corporate investment or accruals is essentially orthogonal to dispensability. These rule-to-characteristic relationships are structural features of the sorting variables themselves; the cross-factor correlation in d_f across halves of +0.958(Panel D) reflects that stability. The 0.421 out-of-sample R^2 in Panel A is therefore not a claim that dispensable stocks remained the same over 45 years, but that each sorting variable's interaction with dispensability did, which is what links pre-period d_f loadings to post-period PreTOM concentration.

IA.7 ETF Evidence

ETFs provide a portfolio-level test using real-money holdings. If dispensability identifies the stocks sold during the cash-demand window, ETFs with greater holdings-weighted dispensability should earn lower PreTOM returns and should exhibit the same T+1 timing shift as factor portfolios.

For each ETF, we compute holdings-weighted dispensability, d_e. ETFs with higher d_e hold more stocks far below their 52-week highs.[1] These ETFs earn lower PreTOM returns. The FF3-controlled slope from

$$\bar{R}_e^{\text{PreTOM}} \;=\; \alpha \;+\; \beta\, d_e \;+\; \beta_{\text{Mkt}}\hat{\beta}_e^{\text{Mkt}} \;+\; \beta_{\text{SMB}}\hat{\beta}_e^{\text{SMB}} \;+\; \beta_{\text{HML}}\hat{\beta}_e^{\text{HML}} \;+\; u_e \tag{1}$$

is -10.59 basis points per day ($t = -2.03$, Table IA.19, Figure IA.2). The Rest slope is small and positive ($+4.33$, $t = +1.80$), consistent with post-window reversal rather than a persistent expected-return relation.

The T+1 reform provides a timing test. High-d_e ETFs exhibit the predicted shift from $\tau{-}4$ to $\tau{-}3$, parallel to the factor-level result. For the 416 ETFs with both pre- and post-reform return windows, we compute the within-month $\tau{-}4$ minus $\tau{-}3$ return differential $y_{e,r}$, separately pre- and post-reform, and estimate

$$y_{e,r} \;=\; \alpha + \beta\, d_e + \gamma\, \text{Post}_r + \delta\,(d_e \times \text{Post}_r) + X_e'\Gamma_r + u_{e,r}, \tag{2}$$

with per-ETF FF3 betas and their interactions with Post_r, SE clustered by ETF. The DiD is $\hat{\delta} = +36.9$ ($t = +2.22$). High-d_e ETFs' $\tau{-}4$ versus $\tau{-}3$ differential rose by an additional $+36.9$ basis points after the reform. The factor-level settlement shift documented in Section 5 thus appears in ETF baskets on the predicted days. The DiD uses a broader cross-section than the level test. Its inclusion requirement is four years of pre-reform $\tau{-}4$ data per ETF, but it does not impose the broad/style/factor/cap-based name filter. The additional ETFs are predominantly sector funds (CRSP objective code `EDS`) plus some country-specific and narrow-thematic funds. The broader universe is appropriate because the DiD is identified within ETF. Across the specification ladder, on the broader 735-ETF sample without the FF3-beta requirement, the raw DiD is $+46.2$ ($t = +4.26$, HC3; $t = +3.82$ ETF-clustered) without FF3 controls; on the matched FF3 sample the controlled DiD is $+36.9$ ($t = +2.22$),

[1] Formally, $d_e = T_e^{-1}\sum_t \sum_i w_{ei,t}\,\delta_{i,t}$, where $w_{ei,t}$ is ETF e's portfolio weight on stock i on day t and $\delta_{i,t}$ is stock i's dispensability score. We restrict the level test to U.S. equity broad, style, factor, and cap-based ETFs (CRSP objective code EDY or EDC), dropping country-specific, narrow-commodity, narrow-biotech, leveraged, and alternative-strategy funds; with TNA $\geq$ \$10M and at least 18 months of PreTOM history, the filter retains 541 of the original 927 ETFs.

the value reported in the main paper. The post-reform ETF window is shorter than the 19-month factor-level window of Section 5 because the ETF analysis additionally requires same-month FF3-beta estimation; the two analyses are not directly comparable on window length.

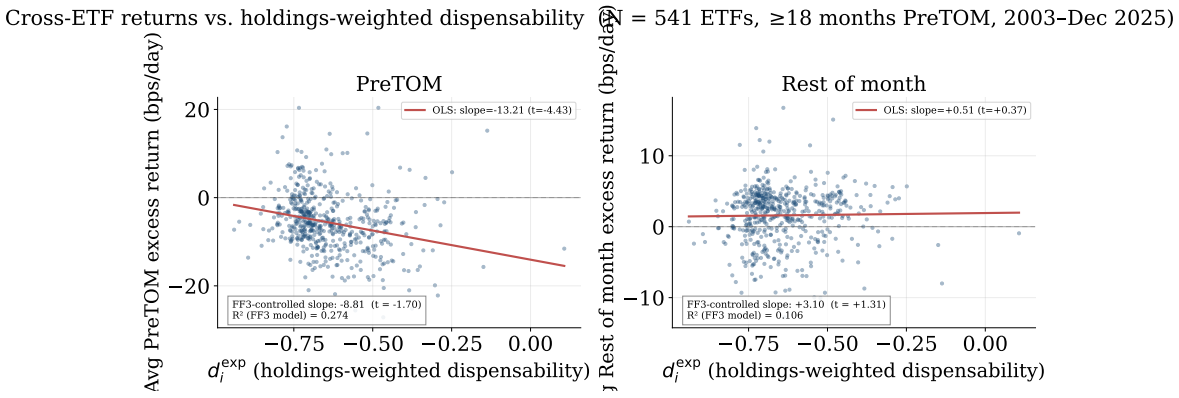

Figure IA.2: 541 U.S. equity broad, style, factor, and cap-based ETFs (TNA $\geq$ \$10M, $\geq$ 18 months of PreTOM history, 2003–Dec 2025). x-axis: holdings-weighted ETF-level dispensability d_e (winsorized 1/99). y-axis: ETF average daily excess return in bps/day (winsorized 0.5/99.5). Left panel: PreTOM ($\tau-9$ to $\tau-4$); right panel: rest of month. The black line is the FF3-controlled slope on d_e, anchored at the cross-section mean. Titles report slope, t-statistic (HC1), and ΔR^2 over the FF3-only model.

The DiD is stable along the specification ladder: it is +92.4 ($t = +6.84$, HC3) pooling raw stock–day observations across 733 ETFs, +46.2 ($t = +3.82$) after collapsing to a per-ETF DiD and clustering by ETF over the broader 735-ETF sample, and +36.9 ($t = +2.22$) in the main specification, which restricts to the 416 ETFs with both return windows and adds FF3-beta $\times$ reform-period interactions.

IA.8 The Stambaugh-Yu-Yuan Mispricing Composite

Data and sign convention. We use the MGMT and PERF clusters from Stambaugh and Yuan (2017), covering 148 of our 153 factors.[2] The MGMT cluster aggregates net stock issues, composite equity issues, accruals, net operating assets, asset growth, and investment. The PERF cluster aggregates momentum, distress, O-score, gross profitability, and return on assets. In SYY, a higher composite score predicts lower returns (overpriced stocks). At the factor level we compute the value-weighted $D_{10}-D_1$ score difference, matching the $\bar{R}^{\text{PreTOM}}_{\text{D10},m} - \bar{R}^{\text{PreTOM}}_{\text{D1},m}$ return convention used throughout the paper. Under this convention,

[2]The SYY composite requires the 11 underlying anomaly variables to be simultaneously available at the stock-month level. A small number of JKP factors (e.g., *ni_inc8q*, *rd5_at*, *div12m_me*) have D1 or D10 portfolios whose overlap with the SYY stock panel is too thin to produce a reliable VW D_1-D_{10} MGMT/PERF gap, leaving 148 factors with usable loadings.

Table IA.19: ETF Evidence: Cross-Section and T+1 Settlement Reform

Panel A: ETF holdings exposure predicts PreTOM returns (541 ETFs)

	PreTOM		Rest of month	
Predictor	Bivariate	FF3-controlled	Bivariate	FF3-controlled
d_e (winsorized 1/99)	-14.74^{***}	-10.59^{**}	$+0.86$	$+4.33^{*}$
	(-5.09)	(-2.03)	$(+0.63)$	$(+1.80)$
β^{Mkt}		-1.34		$+1.96^{**}$
		(-1.04)		$(+2.49)$
β^{SMB}		-0.51		-1.96^{***}
		(-0.39)		(-2.84)
β^{HML}		-7.50^{***}		$+0.30$
		(-5.07)		$(+0.41)$
ΔR^2 from d_e		$+1.23$ pp		$+0.71$ pp
R^2	6.3%	27.6%	0.1%	10.8%
ETFs N	541	541	541	541

Panel B: ETF holdings exposure predicts the T+1 timing shift (416 ETFs)

Predictor	Bivariate	FF3×Post-controlled
d_e (winsorized 1/99)	$+82.0^{***}$	$+36.9^{**}$
	$(+6.46)$	$(+2.22)$
R^2	10.4%	45%
ETFs N	416	416

Panel A notes. Cross-sectional regressions of ETF average daily PreTOM and Rest returns on holdings-weighted dispensability d_e. Sample: 541 U.S. equity broad, style, factor, and cap-based ETFs, 1998–Jan 2025. FF3-controlled columns include full-sample ETF betas on Mkt-RF, SMB, and HML. ETF-universe filters (CRSP objective codes, name screens) are documented in the footnote to this section.

Panel B notes. Cross-sectional regressions of the ETF-level T+1 timing-shift DiD on d_e. Sample: 416 ETFs with both pre- and post-reform return windows around the May 28, 2024 settlement reform. Returns and d_e are winsorized as described in the text of this section. HC1 robust standard errors. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

factors whose long leg is more overpriced (positive MGMT_f or PERF_f) earn lower returns, so the predicted SYY coefficient is negative.

Window-by-window horse race. The main paper reports the cross-factor horse race in Table 4 using bootstrap inference on the canonical $D_{10}-D_1$ specification. The numbers in that table are the canonical record for MGMT and PERF.

Window asymmetry between MGMT and dispensability. The horse race in Table 4 reveals a window asymmetry between the two channels. In PreTOM, d_f alone explains a large share of the cross-factor variation, while MGMT and PERF add only modest incremental fit ($\Delta R^2 \approx 0.08$ for each). In Rest, the ordering reverses: d_f is statistically indistinguishable from zero, while MGMT retains incremental explanatory power for the cross-factor Rest variation (about 8%). PERF behaves similarly to MGMT in the joint regression because the two share several underlying constituents; its incremental fit beyond MGMT is small.

Implications. Our decomposition complements the mispricing literature. After accounting for the PreTOM component, MGMT's Rest explanatory power becomes visible because the calendar-bound friction that contaminates monthly returns is no longer in the dependent variable. Return predictability in the factor zoo splits across two windows: liquidity-demand pressure in PreTOM and corporate-policy characteristics in Rest. The Stambaugh and Yuan composite, by averaging across both, understates the strength of each.

IA.9 Formal Methodology: Characteristic-Loading Framework

This appendix formalizes the characteristic-loading framework used in the main paper. It states the construction in full mathematical detail and reproduces the characteristic-selection logic that identifies distance below the 52-week high as the leading characteristic for cross-factor PreTOM premia. The corresponding informal exposition appears in Sections 3 and 3.2 of the main paper.

From Stock Characteristics to Factor-Level Exposure

We begin by asking which firm characteristic organizes the cross section of factor returns during the PreTOM window. Our goal is not to form a new traded factor. Rather, we use

firm characteristics to measure each anomaly portfolio's exposure to the type of stocks that institutions are likely to liquidate when they face predictable month-end cash demands.

Let $f = 1, \ldots, F$ index the 153 JKP long-short factors, and let $c \in \mathcal{C}$ denote a candidate firm characteristic. For each stock i and month m, we first standardize the characteristic cross-sectionally:

$$z^c_{i,m} \;=\; \frac{c_{i,m} - \bar{c}_m}{\sigma_m(c)}. \tag{3}$$

The standardization makes characteristics comparable across months and across signals measured in different units. Throughout, $c_{i,m}$ is observed at the end of month $m-1$ and used to form the portfolios that earn month-m returns; the index m refers to the return month.

For each factor f, let $S_{f,m}$ and $L_{f,m}$ denote its short and long legs in month m, and let $w^S_{i,f,m}$ and $w^L_{i,f,m}$ denote the corresponding within-leg value weights. We define factor f's loading on characteristic c as the value-weighted short-minus-long difference in characteristic scores:

$$d^c_{f,m} \;=\; \sum_{i \in S_{f,m}} w^S_{i,f,m}\, z^c_{i,m} \;-\; \sum_{i \in L_{f,m}} w^L_{i,f,m}\, z^c_{i,m}. \tag{4}$$

We then average this exposure over time:

$$d^c_f \;=\; \frac{1}{M} \sum_{m=1}^{M} d^c_{f,m}. \tag{5}$$

Thus, d^c_f measures whether the short leg of factor f contains stocks with higher values of characteristic c than its long leg. A high value of d^c_f means that the factor is short the characteristic and therefore benefits if stocks with high c underperform.

For each factor f, we estimate its incremental return during the PreTOM window:

$$R_{f,d} \;=\; \alpha_f + \lambda_f \cdot \mathbf{1}\{d \in \text{PreTOM}\} + \varepsilon_{f,d}, \tag{6}$$

where $R_{f,d}$ is the daily return on factor f and PreTOM denotes trading days $\tau-9$ through $\tau-4$ relative to the last trading day of the month. The coefficient λ_f is the factor's incremental daily return during the PreTOM window.

The central cross-factor test is:

$$\lambda_f \;=\; a + b_c \cdot d^c_f + u_f. \tag{7}$$

If characteristic c organizes exposure to the PreTOM liquidation shock, then factors whose short legs load more heavily on c should earn larger PreTOM returns. The prediction is therefore $b_c > 0$.

This procedure gives each candidate characteristic the same opportunity to explain the cross section of PreTOM factor returns. The object of interest is not the return on a standalone characteristic portfolio. It is the ability of a firm characteristic to explain which existing anomaly portfolios load on the PreTOM shock.

Selecting the Characteristic

We implement this procedure for a broad set of candidate firm characteristics. For each c, we compute d_f^c across the 153 JKP factors and estimate equation (7) in the full sample and in two subperiods, 1980–2002 and 2003–2025.

The results identify distance from the 52-week high as the dominant and most stable organizing characteristic. Let

$$d_{i,m}^{52} \;=\; z_m\left(\frac{H_{i,m}^{252} - P_{i,m}}{H_{i,m}^{252}}\right), \tag{8}$$

where $P_{i,m}$ is the stock price and $H_{i,m}^{252}$ is the stock's 252-trading-day high. The raw distance is largest for stocks trading farthest below their 52-week highs, so higher values mean more dispensable; because standardization is affine-invariant, this equals the negative of the standardized price-to-high ratio, the form used in the replication code. We refer to this variable as distance-from-high dispensability.

The corresponding factor-level loading is

$$d_f^{52} \;=\; \frac{1}{M}\sum_{m=1}^{M}\left[\sum_{i\in S_{f,m}} w_{i,f,m}^{S}\, d_{i,m}^{52} \;-\; \sum_{i\in L_{f,m}} w_{i,f,m}^{L}\, d_{i,m}^{52}\right]. \tag{9}$$

A factor with high d_f^{52} is short stocks that are farther from their 52-week highs than the stocks in its long leg. If month-end cash demand induces institutions to sell positions that are easiest to justify liquidating, then such factors should earn higher returns during PreTOM.

Empirically, d_f^{52} is the characteristic that most strongly and stably explains PreTOM returns in both subperiods. In the full sample, the regression of λ_f on d_f^{52} produces an R^2 of 0.58. The relation is not confined to one era: the R^2 is 0.48 in 1980–2002 and 0.30 in 2003–2025. No other characteristic reaches $R^2 \geq 0.30$ in both subperiods of this λ_f specification.

This stability is important. Some characteristics explain PreTOM returns in the full sample but not consistently across eras. Quality, financial safety, and idiosyncratic volatility have substantial explanatory power in the early sample but weaken sharply after 2003.

Momentum has explanatory power in the late sample but is much weaker in the early sample. Distance from the 52-week high is the only characteristic that remains strong in both periods.

We therefore use d_f^{52} as the baseline measure of factor-level dispensability exposure. The interpretation is an exposure interpretation: distance from the 52-week high measures whether a portfolio is short the stocks most susceptible to predictable month-end liquidation pressure. It need not be a standalone traded factor in every subperiod. The relevant prediction is that existing anomaly portfolios with larger d_f^{52} load more strongly on the PreTOM shock.

The search identifies distance below the 52-week high by the cross-factor R^2 of mean PreTOM returns on each candidate's exposure. To confirm that this fit is specific to the pre-turn window rather than a by-product of searching over 153 characteristics, we hold the characteristic exposures fixed and rerun the full search on 1,000 randomly placed six-day windows, recording the maximum R^2 across candidates in each. Across placebo windows the maximum R^2 averages 0.34 and reaches 0.58 at the 99th percentile, while the true PreTOM window attains 0.72, beyond every placebo draw ($p = 0.001$; Table IA.20).

Table IA.20: Random-Window Placebo Test of the Characteristic Search

Statistic	Value
Actual max R^2 (true PreTOM window)	0.72
Placebo max R^2, mean across windows	0.34
Placebo max R^2, 99th percentile	0.58
Placebo max R^2, maximum across windows	0.67
p-value (actual vs. placebo)	0.001

Notes: 1,000 placebo draws (seed `20260723`). In each draw we select 6 random trading days per month as a placebo window, recompute each factor's mean window long–short return, and rerun the full leave-one-out cross-section of window returns on each of the 153 candidate characteristics' fixed exposures, recording the maximum R^2 across candidates. The characteristic exposures are held fixed; only the return window varies. The actual value is the same maximum R^2 computed on the true PreTOM window (trading days $t-9$ to $t-4$), which reproduces the top of the full-sample search. The p-value is the permutation share of placebo draws at or above the actual value, $(k+1)/(B+1)$. Raw VW long–short decile returns, NYSE breakpoints, 1963–2025.

IA.10 Post-Window Reversal and Dispensability Exposure

Reversal appears in three distinct forms in the paper, and it is worth separating them: (i) the Post-window sign of the dispensability spread itself, which is negative because dispensable stocks recover after month-end (Section 4.1); (ii) the cross-factor relation between average Post-window and average PreTOM returns, examined here; and (iii) the month-by-month question of whether a given month's recovery scales with that month's concession, examined in Section 4.2. The mechanism predicts the first two; the third depends on how quickly liquidity-supplying capital moves, and Section 4.2 finds the recovery is broad and slow rather than concession-scaled.

We regress each factor's average daily Post-window return on its average daily PreTOM return,

$$\bar{R}_f^{\text{Post}} \;=\; \alpha + \beta\, \bar{R}_f^{\text{PreTOM}} + \varepsilon_f, \tag{10}$$

across all factors, one observation per factor. Pooled across all factors the slope is -0.07 and insignificant. The question of interest is whether the slope depends on a factor's exposure to the liquidity-demand component, which we test by adding a dispensability interaction,

$$\bar{R}_f^{\text{Post}} \;=\; \alpha + \beta\, \bar{R}_f^{\text{PreTOM}} + \theta\, \tilde{d}_f + \gamma \left(\bar{R}_f^{\text{PreTOM}} \times \tilde{d}_f\right) + \varepsilon_f, \qquad \tilde{d}_f \equiv |d_f| - \overline{|d|}, \tag{11}$$

where $\tilde{d}_f$ is the within-sample-centered absolute exposure, so β is the Post-on-PreTOM slope at average exposure and the marginal slope at exposure $|d_f|$ is $\beta + \gamma\, \tilde{d}_f$. The level term $\theta\, \tilde{d}_f$ allows Post-window returns to differ with exposure directly; the interaction γ is the object of interest.

The slope falls significantly with exposure ($\gamma = -0.35$; month-block bootstrap $t = -2.20$; OLS $t = -3.71$), and the slope at average exposure is essentially zero ($\beta = +0.02$). Where dispensability exposure is low there is no transient component, so the only cross-window commonality is each factor's standing premium, earned on all days; the Post and PreTOM returns share it and the slope is positive (continuation). Where exposure is high the PreTOM return is dominated by the transient liquidity-demand concession, which unwinds after month-end, so a larger PreTOM return is followed by a smaller Post return and the slope turns negative (reversal). The fitted marginal slope is positive at low $|d_f|$ and crosses into negative territory at high $|d_f|$; the pooled slope is near zero precisely because these regions cancel.

Terciles of $|d_f|$ show the same pattern (Table IA.21, Panel B). The slope is $+0.55$ ($t = +3.71$) at low exposure and -0.25 ($t = -1.18$) at high exposure. Binning into $\approx$ 48-

factor groups is underpowered, so the high-tercile reversal has the predicted sign but is not significant under the bootstrap; the continuous interaction is the formal test, and the terciles are the visualization. The interaction is stronger under a scale-free rank specification, both legs are individually significant in a two-piece below/above-median fit, and no single factor drives it (leave-one-factor-out keeps γ within $[-0.41, -0.27]$).[3]

[3]Value is the one influential theme. Dropping it roughly halves γ, to an insignificant -0.22 ($t = -1.34$), consistent with value factors, which are long dispensable stocks ($d_f < 0$), sitting at the high-exposure, reversing end of the spectrum. Dropping Momentum, Low Risk, or Profitability leaves the interaction significant.

Table IA.21: Post-Window Reversal of the PreTOM Return

Panel A: Continuous interaction across all 144 factors

$\bar{R}_f^{\text{Post}} = \alpha + \beta\,\bar{R}_f^{\text{Pre}} + \theta\,\tilde{d}_f + \gamma\left(\bar{R}_f^{\text{Pre}} \times \tilde{d}_f\right) + e_f, \quad \tilde{d}_f \equiv |d_f| - \overline{|d|}$

	β (slope at avg $\lvert d_f\rvert$)	θ ($\tilde{d}_f$)	γ (interaction)
Coefficient	+0.02	+2.19	-0.35
t (month-block bootstrap)			-2.20

Panel B: By $|d_f|$ tercile

Subset	N	β (t)	R^2
All factors	144	-0.07 (-0.44)	0.013
Q1: low $\lvert d_f\rvert$ (persistence benchmark)	48	+0.55 (+3.71)	0.348
Q2: middle $\lvert d_f\rvert$	48	+0.08 (+0.47)	0.012
Q3: high $\lvert d_f\rvert$ (reversal)	48	-0.25 (-1.18)	0.190

Notes: Each observation is a factor; the dependent variable is the factor's average daily Post-window return. Panel A regresses it on the average daily PreTOM return, the absolute dispensability exposure $|d_f|$, and their interaction, across the 144 JKP factors with PreTOM, Post-window, and dispensability-exposure statistics available (Section 2). The interaction $\gamma < 0$ means the PreTOM-to-Post slope falls as dispensability exposure rises; the transient PreTOM concession unwinds in the Post window, and it does so more for factors more exposed to dispensable stocks. Panel B reports the same Post-on-PreTOM regression (equation 10) within terciles of $|d_f|$, with d_f defining the subsamples (not a regressor); it shows the same pattern (persistence at low $|d_f|$, reversal at high), but binning into ≈48-factor terciles discards the power that the pooled interaction retains. Slopes and R^2 are full-sample OLS; t-statistics are a month-block bootstrap ($B = 5{,}000$, resampling whole calendar months and recomputing each factor's window means within the draw), with $|d_f|$ held fixed at the full-sample sort.

IA.11 Cross-Factor Horse Race: Candidate Construction

This appendix documents the construction of the eight candidate predictors used in the cross-factor horse race (Section 3.4, Table 4). Each candidate is summarized by a single per-factor scalar that we regress against the dispensability exposure d_f and the average daily long-short return $\bar{R}_f$ in PreTOM and Rest. All candidates are cross-sectionally standardized (mean 0, standard deviation 1) within the regression so that coefficients are comparable across rows. Throughout, "value-weighted $D1-D10$ difference" weights each stock by its month-end market capitalization, contemporaneous with the underlying characteristic.

S.1. Mechanical-overlap and risk-loading candidates

UMD overlap. For each factor f we compute the value-weighted holdings overlap between f's deciles and the Jegadeesh and Titman (1993) 12–2 momentum sort. Overlap is defined as the time-average market-cap-weighted correlation between the indicator vector of stocks in f's extreme decile and the indicator vector of stocks in UMD's matching extreme decile. Construction uses portfolio composition rather than realized returns, so the test does not condition on the dependent variable. Sample: 1963–2025 monthly holdings.

Market beta (d_f^{β}). The factor-level market-beta loading is the time-averaged D10−D1 difference of stock-level 60-month rolling market beta (JKP `beta_60m`) within each factor's NYSE-breakpoint deciles. Sample: 1963–2025.

S.2. Liquidity- and friction-based candidates

Pastor-Stambaugh aggregate liquidity (β_f^L). For each factor we estimate a monthly time-series regression of the long-short return on the Pástor and Stambaugh (2003) non-traded liquidity-innovation series (eq. 8 in the original paper), controlling for the Fama-French three factors with Newey-West HAC standard errors at four lags:

$$R_{f,m}^{LS} = a_f + b_f^{\text{Mkt}}\text{MktRF}_m + b_f^{\text{SMB}}\text{SMB}_m + b_f^{\text{HML}}\text{HML}_m + \beta_f^L\text{LIQ}_m + \varepsilon_{f,m}.$$

β_f^L is the FF3-controlled slope. Sample: 1962–2024 (PS liquidity-data window).

Amihud illiquidity (β_f^A). The factor-level Amihud loading is the time-averaged D1−D10 difference of stock-level six-month rolling Amihud illiquidity (JKP `ami_126d`) within each factor's NYSE-breakpoint deciles. This factor-level exposure is distinct from the market-wide time-series Amihud illiquidity control in the state-dependence analysis of Section 4.3. Sample: 1980–2024 (Amihud panel-availability window).

Shorting cost (SIRIO). We follow Drechsler and Drechsler (2021) in proxying shorting cost by the ratio of short interest to institutional ownership. SIRIO is computed at the stock-month level from Compustat short interest and 13F institutional holdings, then aggregated to the factor level as the time-averaged value-weighted D1−D10 difference within each factor's NYSE-breakpoint deciles. Sample: 2003–2025 (matched window for both the SIRIO panel and the corresponding $\bar{R}_f$ in the horse-race regression).

S.3. Belief- and sentiment-based candidates

Subjective Belief Factor (β_f^{SBF}). The SBF series is the realized one-quarter-ahead state variable s_{t+1} from Cui, De la O, and Myers (2025), sampled quarterly 1982Q4–2022Q2 and

provided directly by the authors. We forward-fill the quarterly series to daily ($\text{SBF}_d = s_{Q(d)}$ for days d in realization quarter Q) and estimate per-factor daily time-series regressions with FF3 controls and Newey-West HAC standard errors at 60 lags:

$$r_{f,d} = a_f + b_f^{\text{Mkt}}\text{MktRF}_d + b_f^{\text{SMB}}\text{SMB}_d + b_f^{\text{HML}}\text{HML}_d + \beta_f^{\text{SBF}}\text{SBF}_d + \varepsilon_{f,d}.$$

β_f^{SBF} is the SBF loading per factor. Sample: 1982Q4–2022Q2.

MGMT and PERF mispricing factors. We follow Stambaugh and Yuan (2017) (Section IA.8) and use their MGMT and PERF composites, each constructed as an average of percentile ranks across constituent anomaly signals. The factor-level loadings MGMT_f and PERF_f are the time-averaged value-weighted D1−D10 differences of stock-level MGMT and PERF scores within each factor's NYSE-breakpoint deciles. Sample: 1963–2025.

S.4. Sample-matching for the horse-race regression

For each candidate, both the dependent variable $\bar{R}_f$ and the candidate X_f are recomputed on a sample window matched to the candidate's data availability: 1980–2024 for Amihud, 2003–2025 for SIRIO, 1982Q4–2022Q2 for SBF, and 1963–2025 for the remaining five candidates. Matched-sample d_f-alone R^2 values appear in column 3 of Table 4; the ΔR^2 column reports the joint-minus-d_f-alone difference, which isolates each candidate's incremental cross-sectional content beyond what dispensability captures on the same sample.

IA.12 Short-Leg vs Long-Leg Decomposition

We decompose each factor's PreTOM premium into the contribution from its short leg (the lower-prc decile) and its long leg (the higher-prc decile), each signed so that "factor wins" is positive ($y_f^{\text{short}} = -\bar{r}_f^{\text{short,Pre}}$ and $y_f^{\text{long}} = +\bar{r}_f^{\text{long,Pre}}$ in daily basis points). Each leg is then regressed on the absolute dispensability gap $|d_f|$ across all 153 factors, with theme-clustered standard errors:

Dependent (bps/day)	$\hat{b}$	t	R^2	Mean
y_f^{short} in PreTOM	$+\mathbf{2.77}$	$+\mathbf{10.23}$	0.371	+5.97
y_f^{long} in PreTOM	+1.26	+5.26	0.237	−3.64
y_f^{short} in Rest	+0.25	+0.84	0.011	—
y_f^{long} in Rest	−0.43	−0.76	0.032	—

Of the cross-sectional spread, $69\% = 2.77/(2.77 + 1.26)$ comes from the short leg and 31% from the long leg; both vanish in Rest.

IA.13 Theme-Level PreTOM versus Rest, Daily Long–Short Returns

Figure IA.3 summarizes theme-level concentration of factor returns in the PreTOM window.

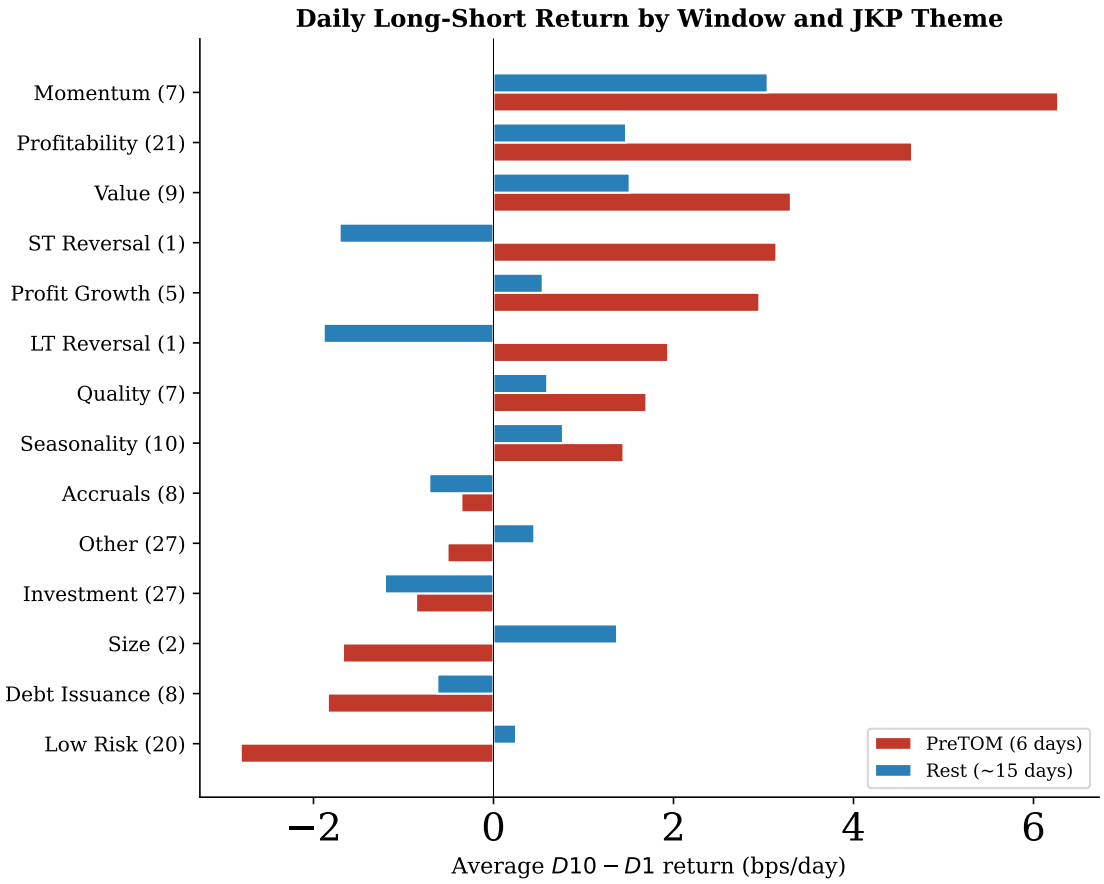

Figure IA.3: Average daily long–short return (bps/day) during the PreTOM window (red) and the rest of the month (blue) for each JKP theme. Returns are the raw, unsigned $D10-D1$ spread (no per-factor intended-sign flip), so each theme falls on the side of zero implied by its mechanical decile ordering rather than by the anomaly's traded direction. Number of factors per theme in parentheses.

IA.14 Stock-Level Fama-MacBeth Specification

This appendix gives the two-stage Fama-MacBeth specification underlying the stock-level results reported in Section 6 of the main paper. In the first stage, for each trading day t, we regress stock-level market-adjusted returns on dispensability-quintile dummies and three Fama-French controls (lagged 60-month market beta, log market capitalization, and book-to-market):

$$r_{i,t} - r_t^m \;=\; a_t \;+\; \sum_{q=2}^{5} b_{q,t}\,\mathbf{1}\{Q_{i,m-1} = q\} \;+\; \gamma_t' X_{i,m-1} \;+\; u_{i,t}, \tag{12}$$

$Q_{i,m-1} \in \{1, \ldots, 5\}$ is stock i's NYSE-breakpoint quintile of the dispensability score $D_{i,m-1}$ at month-end $m-1$, with Q_1 (least dispensable, near-52-week-high) omitted. $X_{i,m-1} = (\beta_{i,m-1}, \log ME_{i,m-1}, BM_{i,m-1})$ is the control vector.

The second stage decomposes each daily slope into Rest and PreTOM components:

$$b_{q,t} \;=\; c_0^q + c_1^q\,\mathrm{PreTOM}_t + v_{q,t}, \quad q \in \{2, \ldots, 5\}, \tag{13}$$

with Newey-West standard errors at five lags. $\hat{c}_1^5$ is the Q5−Q1 PreTOM-minus-Rest differential. The quintile coefficients are $\hat{c}_1^2 = -2.09$, $\hat{c}_1^3 = -4.25$, $\hat{c}_1^4 = -3.57$, and $\hat{c}_1^5 = -6.14$ in bps/day, all with $|t| > 2.8$ (Table IA.22). Stocks far from their 52-week highs underperform stocks near them by roughly six basis points per day during the six PreTOM days after controlling for the three Fama-French firm characteristics.

Table IA.22: Stock-Level Fama–MacBeth: Q2−Q1 through Q5−Q1 Dispensability Spreads, PreTOM vs Rest

	Value-weighted		Equal-weighted	
$\hat{c}_1^q$ (PreTOM − Rest)	Baseline	+FF3	Baseline	+FF3
Q2	−2.81	−2.09	−1.50	−1.16
	(−3.70)	(−3.20)	(−3.29)	(−2.85)
Q3	−4.35	−4.25	−2.35	−2.21
	(−4.04)	(−4.76)	(−3.44)	(−3.70)
Q4	−4.72	−3.57	−3.38	−3.05
	(−3.23)	(−2.99)	(−3.60)	(−3.74)
Q5	−7.91	−6.14	−7.11	−7.16
	(−3.65)	(−3.55)	(−4.61)	(−5.56)
$\beta_{i,m-1}$		−3.58		−2.98
		(−2.73)		(−3.75)
$\log ME_{i,m-1}$		+0.20		−0.79
		(+0.80)		(−2.97)
$BM_{i,m-1}$		−0.90		−0.31
		(−1.62)		(−0.75)
Stock-day obs	57,887,080 (full CRSP, non-missing D)			
Trading days	15,834 (4,530 PreTOM, 11,304 Rest)			
Sample	February 1963 – December 2025			

Notes: Each cell reports the PreTOM-versus-Rest differential $\hat{c}_1^q$ from the two-stage Fama–MacBeth procedure. Stage 1 runs a daily cross-sectional regression of stock market-adjusted returns on D-quintile dummies (Q1, least dispensable, omitted) with optional FF3 firm controls (60-month market beta, log market cap, book-to-market); quintiles use NYSE breakpoints. Stage 2 regresses each daily quintile slope on a PreTOM dummy, Newey–West 5 lags; $\hat{c}_1^q$ is the coefficient on PreTOM. VW columns weight stage 1 by lagged market cap. 1963–Dec 2025.

IA.15 Additional Tables Referenced in the Main Paper

Tables in this section are referenced in the main paper but moved here to keep the main text focused on load-bearing results.

Dispensability Mapping for Canonical Anomalies

Table IA.23: Dispensability Mapping for Canonical Anomalies (Raw Spreads R_f)

Factor	Dispensable (short) leg	sign d_f	PreTOM $D10-D1$
Momentum (UMD)	$D1$ (past losers)	+	+10.4
Quality (QMJ)	$D1$ (junk)	+	+6.2
Gross profitability	$D1$ (unprofitable)	+	+3.7
MGMT (Stambaugh–Yuan)	$D1$	+	+3.5
Low-beta (BAB)	$D10$ (high beta)	−	−5.0
Value (B/M)	$D10$ (high B/M)	−	−2.2

Notes: For each factor we report its dispensable (short) leg, the sign of its short-minus-long dispensability exposure d_f, and its raw $D10-D1$ PreTOM return (average daily, bps, 1963–2025). A positive d_f means the factor is short the dispensable side, so month-end selling raises its $D10-D1$ PreTOM return; a negative d_f means it is long the dispensable side, so the same selling lowers it. The sign of d_f and the sign of the PreTOM return coincide by construction.

Y.1. Within-Leg Primitive Horse Race

Table 10 reports the pooled within-leg Q5−Q1 spread using distance below the 52-week high as the sub-sorting variable. Table IA.24 repeats the same within-leg sub-sort with eight alternative candidate characteristics against the same incumbent, distance below the 52-week high (`prc_highprc_252d`). Five are the data-selected top rivals from the Section 3.2 cross-sectional search ranked by cross-sectional R^2 on λ_f (F-score, QMJ, the Stambaugh–Yuan PERF mispricing composite, 21-day realized volatility, and the 21-day Hou–Xue–Zhang idiosyncratic volatility); the remaining three are further characteristics emphasized in the related literature as proxies for stock quality, profitability, and idiosyncratic risk (QMJ-safety, operating profitability, and the 21-day CAPM idiosyncratic volatility).

For each candidate, we apply the within-leg sub-sort: inside each of the six short legs of Table 10 we sub-sort stocks into five within-leg quintiles by the candidate's within-month z-score (Q1 the lowest z, Q5 the highest), then compute the daily value-weighted Q1−Q5 spread and stack the six daily series. The pooled regression is $\text{spread}_t = \alpha + \beta \cdot \mathbf{1}[\text{PreTOM}_t] + \varepsilon_t$ with standard errors clustered by short-leg factor. The PreTOM column reports $\alpha + \beta$; the

Rest column α; the DiD column β, the PreTOM-minus-Rest differential that isolates the calendar-priced component of the within-leg sort. Sign conventions differ across rivals: for the 52-week-high characteristic (the raw JKP price-to-high ratio), lower z marks more dispensable stocks (so Q1−Q5 is negative when dispensable stocks underperform); for realized and idiosyncratic volatility, higher z marks more dispensable stocks (so Q1−Q5 has the opposite sign). The DiD is the right discriminator across these sign conventions: characteristics whose within-leg sort is calendar-invariant produce similar PreTOM and Rest spreads (DiD near zero), while the dispensability axis produces a PreTOM-specific differential.

Distance below the 52-week high has DiD $= -7.83$ bps/day ($t = -11.8$), more than 2.5× the absolute DiD of any other non-confounded candidate. The next-strongest same-signed rival is operating profitability at -2.48 ($t = -2.8$); the next-strongest of any sign is 21-day realized volatility at $+3.10$ ($t = +2.7$, sign reversed for the volatility cluster as noted above); the remaining non-confounded rivals all have $|t_{\text{DiD}}| < 2$. Within-leg sub-sorts on rival characteristics produce substantial pooled spreads (10−13 bps/day in absolute value), but symmetrically across the two windows. Only the 52-week-high axis produces a within-leg spread that concentrates in PreTOM, consistent with a calendar-priced dispensability effect rather than generic quality or risk dispersion within the short legs. Two cells are flagged as confounded (within-leg z-score standard deviation below 0.5): QMJ inside the QMJ short leg (sub-sorting on QMJ within QMJ junk is near-trivial) and QMJ-safety inside the QMJ short leg. The flag is descriptive; even taking the flagged cells at face value, neither matches the 52-week-high DiD.

Table IA.24: Within-Leg Horse Race: Candidate Sub-Sorting Characteristics

Sort variable (Q1−Q5)	PreTOM (bps/d)	t	Rest (bps/d)	t	DiD (bps/d)	t
`prc_highprc_252d` (incumbent)	-13.22	-20.45	-5.39	-6.24	-7.83	-11.58
`rvol_21d`	+12.95	+35.18	+9.85	+8.52	+3.10	+2.74
`op_at`	-12.44	-29.70	-9.96	-9.28	-2.48	-2.86
`ivol_capm_21d`	+11.69	+14.54	+9.68	+8.59	+2.01	+3.28
`mispricing_perf`	-10.03	-8.74	-9.16	-16.28	-0.87	-0.67
`ivol_hxz4_21d`	+9.87	+20.55	+10.04	+10.13	-0.17	-0.20
`f_score`	-5.41	-4.77	-4.07	-15.19	-1.34	-1.29
`qmj`*	-5.04	-4.78	-3.14	-4.05	-1.89	-1.69
`qmj_safety`*	-4.42	-4.28	-1.24	-1.05	-3.18	-2.34

Notes: Pooled within-leg Q1−Q5 daily VW spreads (basis points per day, in excess of the risk-free rate) across six short-leg samples: `ret_12_1` D1, `gp_at` D1, `op_at` D1, `o_score` D10, `ni_me` D1, and `qmj` D1. Inside each factor's short-leg decile we sub-sort stocks each month into five quintiles by the sort variable's within-month z-score; the spread is daily VW return of Q1 minus VW return of Q5. Pooled regression: $\text{spread}_t = \alpha + \beta \cdot \mathbf{1}[\text{PreTOM}_t] + \varepsilon_t$ with standard errors clustered by short-leg factor across the six legs. Rows ranked by absolute pooled PreTOM spread; the incumbent 52-week-high characteristic is listed first. The first five rivals are the data-selected top-five by cross-sectional R^2 on λ_f; the last three are reviewer-named quality/risk proxies. *One or more short-leg cells flagged as confounded (within-leg z-score standard deviation below 0.5, indicating that the rival characteristic carries little independent within-leg variation; the spread for those cells is mechanically attenuated).

IA.16 Robustness to Within-Theme Factor Redundancy

The 153 JKP characteristics cluster into 14 themes (Profitability, Momentum, Value, and so on); characteristics within a theme are mechanically related, and a critic could argue that this within-theme redundancy inflates the cross-sectional R^2 of d_f on λ_f. Four diagnostics compare redundancy in PreTOM and Rest. The main result, in Table IA.25, is that the eigenstructure of the 152-factor daily cross-section is essentially identical in PreTOM and Rest, so any redundancy present in PreTOM is also present in Rest. The order-of-magnitude gap in d_f's cross-sectional R^2 across the two windows is therefore not a redundancy artifact, since the redundancy is symmetric. The remaining three tables corroborate the finding under correlation-pruning, a multiple-testing bootstrap over 153 candidate characteristics, and JKP-theme equal weighting. All inference uses a month-block bootstrap with $B = 5{,}000$ draws; both $\bar{R}_f$ and d_f are recomputed inside each draw.

Z.1. Eigenstructure window-invariance (PCA)

We form the daily long–short return matrix $R \in \mathbb{R}^{T \times F}$ for the JKP factor cross-section (142 factors with at least 95% daily coverage; `prc_highprc_252d` excluded as the self-loading factor) and compute the eigendecomposition of $R^\top R$ separately on PreTOM days, Rest days, and the full sample. The first principal component's variance share and its correlation with d_f are indistinguishable across windows (Table IA.25). Higher-order components show similarly small PreTOM-Rest differences. The dispensability exposure d_f spans the top eigenspace (PC1, PC2, and PC4 together account for 94% of d_f's directional variance) but is not concentrated on any single component. Under a covariance decomposition the dispensability axis loads even more heavily on the leading component; the point relevant to redundancy is that this eigenstructure is essentially the same in PreTOM and Rest, so any within-theme redundancy is symmetric across the two windows.

Z.1a. Frontier counterfactual

Table IA.27 reports the counterfactual discussed in Section 3.8: annualized Sharpe ratios of the zoo's portfolios and of its tangency frontier, computed from full-sample moments and from Rest-day moments alone.

Table IA.25: Robustness Tier 1.1: PCA Decomposition of the Daily Factor-Return Matrix

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Panel A: Variance explained (individual, covariance PCA)										
Full	28.2%	19.3%	6.3%	4.6%	3.4%	2.1%	1.8%	1.6%	1.3%	1.2%
Pretom	27.9%	19.5%	6.3%	4.9%	3.5%	2.1%	1.8%	1.5%	1.3%	1.3%
Rest	28.4%	19.2%	6.4%	4.6%	3.3%	2.1%	1.8%	1.6%	1.3%	1.2%
Panel B: Cumulative variance explained										
Full	28.2%	47.6%	53.9%	58.5%	61.9%	64.0%	65.8%	67.3%	68.7%	69.9%
Pretom	27.9%	47.4%	53.7%	58.6%	62.1%	64.2%	66.0%	67.4%	68.7%	70.0%
Rest	28.4%	47.7%	54.0%	58.6%	61.9%	64.0%	65.8%	67.4%	68.8%	70.0%
Panel C: $\lvert\rho\rvert$ between d_f and PCk factor-space loadings (PC1–PC5)										
Full	+0.633	+0.597	+0.096	+0.444	+0.056	—	—	—	—	—
Pretom	+0.619	+0.602	+0.050	+0.456	+0.027	—	—	—	—	—
Rest	+0.637	+0.596	+0.107	+0.434	+0.090	—	—	—	—	—

Notes: Principal component analysis of the daily long–short return matrix of the JKP factor universe (142 factors with at least 95% daily coverage; `prc_highprc_252d` excluded as the self-loading factor), applied separately to the full sample (1963–2025), to PreTOM days ($\tau-9$ through $\tau-4$), and to Rest days. Covariance PCA (demean only, no per-factor standardization) is used. Panel A reports the share of total variance explained by each principal component; Panel B reports the cumulative share. Panel C reports the Pearson correlation (in absolute value, sign-aligned) between the dispensability exposure d_f and the factor-space loadings (right-singular vectors) of PC1–PC5. Effective-rank estimators (Bai–Ng IC_2, Onatski, Ahn–Horenstein) are reported in Table IA.26.

Table IA.26: Robustness Tier 1.3: Effective Number of Factors

Estimator	$\hat{K}$	Notes
Bai–Ng (2002) IC_2, $k_{\max} = 30$	10	Primary integer estimator
Onatski (2010) edge-eigenvalue test	2	Converged
Trace ratio $\mathrm{tr}(\Sigma)^2/\mathrm{tr}(\Sigma^2)$	12.84	Continuous (stable rank)
Ahn–Horenstein (2013) ER, $k_{\max} = 30$	2	Cross-check

Notes: Estimators of the effective number of latent factors in the daily long–short return matrix of the 152-factor JKP universe (`prc_highprc_252d` excluded), $T = 15834$ trading days, 1963–2025. Eigenvalues are computed from the pairwise-complete daily LS correlation matrix. Bai–Ng IC_2 uses $V(k) = N^{-1}\sum_{j>k}\lambda_j$ and $g(N,T) = (N+T)/(NT)\log\min(N,T)$. Onatski (2010) iteratively calibrates the edge gap δ from the bulk slope of eigenvalues and selects the largest $i \le k_{\max}$ with $\lambda_i - \lambda_{i+1} \ge \delta$. The trace ratio reports the continuous "stable rank" $\mathrm{tr}(\Sigma)^2/\mathrm{tr}(\Sigma^2)$, which equals k for an identity-k matrix. The Ahn–Horenstein (2013) eigenvalue ratio is reported as a cross-check.

Z.2. Correlation pruning

We prune the 152-factor cross-section by extracting connected components of the pairwise correlation graph above thresholds $\tau \in \{0.85, 0.90, 0.95\}$ and reporting two specifications: a *medoid* dedup (keep the component centroid) and a *weighted* specification (each factor weighted by $1/|\text{component}|$). The PreTOM slope and bootstrap t-statistic are stable across the pruning ladder (Table IA.28); the most aggressive pruning ($\tau = 0.85$, 123 components) drops PreTOM R^2 from 72.4% to 69.7% while t remains above +4.

Z.3. Multiple-testing bootstrap over 153 candidates

To rule out that the main R^2 is a multiple-testing artifact of selecting d_f from 153 candidate characteristics, we re-run the search inside a month-block bootstrap. In each draw, $\bar{R}^{\text{Pre}}_f$ and the 153×153 matrix of factor-level loadings CE^c_f are recomputed over the same resampled months, and 153 cross-sectional R^2s are computed (one per candidate). Distance below the 52-week high (`prc_highprc_252d`) is the top-1 candidate in 80.6% of bootstrap draws and is in the top-five in 91.4% (Table IA.29). The observed top R^2 of 71.3% and gap to second place of 23.3 pp lie inside the bootstrap distribution's central mass (empirical right-tail p-values $\Pr(R^2_{\text{boot}} \geq 0.713) = 0.12$ and $\Pr(\text{gap}_{\text{boot}} \geq 23.3\text{pp}) = 0.16$), so the observed values are typical rather than tail draws. The 52-week-high characteristic is the leading single predictor in the cross-section.

Z.4. JKP-theme equal weighting

Finally, we re-weight the cross-sectional regression so that each JKP theme contributes equally regardless of its factor count: each factor receives weight $1/N_{\text{theme}(f)}$, with the weights normalized so $\bar{w} = 1$. PreTOM R^2 moves from 72.4% to 77.3% and t remains +4.17 (Table IA.30); Rest R^2 collapses to 0.7%. The result holds in a specification in which large themes such as Investment ($N = 27$) and Profitability ($N = 21$) do not dominate the regression through their factor counts.

Z.5. Theme-collapsed regression

A stronger version of the same exercise collapses the cross-section outright: each of the 14 JKP themes becomes a single observation, the equal-weighted mean across its factors of the intended-signed window return and of the dispensability exposure d_f. Rerunning the cross-sectional regression on the 14 theme observations gives a PreTOM R^2 of 86% (slope 4.82, month-block bootstrap $t = +3.99$) against a Rest R^2 of 1.8% ($t = +0.33$; Table IA.31).

Table IA.27: Sharpe Ratios and Window Shares of Three Zoo Portfolios

	Annualized Sharpe ratio			Window share	Pre−Rest SR
	Full	Rest only	PreTOM only	of mean	diff. [95% CI]
Equal-weighted zoo	0.86	0.65	1.40	46%	+0.74 [+0.17, +1.34]
Dispensability-weighted	0.46	0.18	1.18	72%	+1.00 [+0.37, +1.65]
Leading PC, Rest-estimated	0.24	0.06	0.72	83%	+0.66 [+0.05, +1.28]

Notes: Annualized Sharpe ratios of portfolios of the 142 JKP factors with at least 95% daily coverage over 1963–2025, of the 152 factors used elsewhere (the 52-week-high factor is excluded). Sharpe ratios are per-trading-day moments annualized by $\sqrt{252}$; the Rest-only and PreTOM-only columns use the moments of the indicated days alone, so each is the hypothetical calendar in which every trading day is of that type. Under this coverage screen the equal-weighted window share of the mean is 46%, compared with the 47% headline share on the balanced sample of Table 2. The table reports the equal-weighted portfolio of intended-signed factor returns; the portfolio weighting raw $D10 - D1$ returns by the raw dispensability gap $\widetilde{d}_f$; and the leading principal component of the Rest-day covariance matrix, whose loadings correlate 1.000 with the full-sample leading eigenvector and 0.64 with the dispensability exposure. Each portfolio is oriented so its full-sample mean is positive, and the window share of the mean is the share of its cumulative return earned on PreTOM days. The last column reports the month-block bootstrap mean and 95% confidence interval (B = 5,000 draws, seed 20260810) of the PreTOM-minus-Rest annualized Sharpe difference.

Table IA.28: Robustness Tier 1.2: Correlation-Threshold Pruning of the Factor Cross-Section

		Counts		PreTOM			Rest		
τ	Specification	Components	N used	$\hat{\beta}$	R^2	t	$\hat{\beta}$	R^2	t
—	baseline (no pruning)	152	152	+5.05	72.4%	+4.06	+0.34	1.0%	+0.39
0.85	dedup (medoid)	123	123	+5.58	69.7%	+4.26	+0.20	0.3%	+0.23
0.85	weighted (1/comp size)	123	152	+5.42	67.5%	+4.19	+0.24	0.3%	+0.27
0.90	dedup (medoid)	134	134	+5.26	72.0%	+4.13	+0.36	1.0%	+0.41
0.90	weighted (1/comp size)	134	152	+5.16	70.3%	+3.99	+0.26	0.5%	+0.30
0.95	dedup (medoid)	145	145	+5.07	72.8%	+4.08	+0.41	1.5%	+0.49
0.95	weighted (1/comp size)	145	152	+5.06	72.7%	+4.10	+0.41	1.5%	+0.48

Notes: Correlation-threshold pruning of the 152-factor JKP cross-section. For each threshold $\tau \in \{0.85, 0.90, 0.95\}$, we form an undirected graph on the factors with edges where the pairwise-complete daily long–short return correlation has $|\rho| > \tau$, and identify connected components ("empirical families"). The *dedup (medoid)* specification keeps one factor per component, where the medoid is the factor with the highest average $|\rho|$ to other members (alphabetical tie-break). The *weighted* specification keeps all 152 factors and weights each by 1/component size. Slope $\hat{\beta}$ and R^2 are from the cross-sectional regression of $\bar{R}_f^{\text{Pre}}$ (or $\bar{R}_f^{\text{Rest}}$) on d_f. t uses a month-block bootstrap with $B = 5{,}000$ draws (seed 20260513), resampling calendar months with replacement and recomputing $\bar{R}_f$ inside each draw. In the spirit of Green, Hand, and Zhang (2017) and Feng, Giglio, and Xiu (2020).

Table IA.29: Robustness Tier 3.1: Multiple-Testing Bootstrap over 153 Candidates

Quantity	Observed	Bootstrap distribution
Top cross-sectional R^2	71.3%	mean 61.2%, [2.5%, 97.5%] = [38.4%, 75.9%
Gap to second-place candidate	23.3%	mean 13.0%, [2.5%, 97.5%] = [0.3%, 31.5%
P(boot top $R^2 \geq$ observed)	0.1216	
P(boot gap $\geq$ observed)	0.1646	
Frequency `prc_highprc_252d` ranks in top-1 in bootstrap	80.6%	
Frequency `prc_highprc_252d` ranks in top-3 in bootstrap	88.3%	
Frequency `prc_highprc_252d` ranks in top-5 in bootstrap	91.4%	

Notes: In each of $B = 5{,}000$ month-block bootstrap draws, calendar months are sampled with replacement, and both $\bar{R}_f^{\text{Pre}}$ (153-dim) and CE_f^c (the 153×153 matrix of D1−D10 dispensability exposures for each factor f and candidate characteristic c) are recomputed over the resampled months. The cross-sectional regression of $\bar{R}_f^{\text{Pre}}$ on CE_f^c is run for each of the 153 candidates; the maximum R^2 and the gap to second place are recorded per draw. The observed top-place winner is `prc_highprc_252d` (the 52-week-high characteristic), with $R^2 = 71.3\%$ and a 23.3 pp gap to second-place `f_score`. Empirical p-values are right-tail probabilities (observed values $\geq$ the bootstrap distribution).

Table IA.30: Robustness Tier 3.2: JKP-Theme-Weighted Regression

Specification	PreTOM			Rest		
	$\hat{\beta}$	R^2	t	$\hat{\beta}$	R^2	t
unweighted (baseline)	+5.053	72.4%	+4.06	+0.338	1.0%	+0.39
JKP-theme weighted ($1/N_{\text{theme}}$)	+4.994	77.3%	+4.17	+0.224	0.7%	+0.28

Notes: Cross-sectional regression of $\bar{R}_f^{\text{Pre}}$ (or $\bar{R}_f^{\text{Rest}}$) on the dispensability exposure d_f, with each factor weighted by $w_f = 1/N_{\text{theme}(f)}$ using the JKP 14-theme classification, normalized so $\overline{w} = 1$. Each JKP theme contributes equally regardless of its factor count. The unweighted baseline matches the headline cross-section in the main text. Inference: month-block bootstrap with $B = 5{,}000$ draws.

Whatever redundancy exists within a theme is averaged away entirely before the regression is run.

IA.17 Factor Momentum: Calendar-Component Derivation

This section derives the calendar-anchored component of factor momentum asserted in Section 7.3 of the main paper. Let $g_m \equiv \beta\,\mathrm{LDS}_m^{\mathrm{PreTOM}}$ be the priced liquidity-demand shock, and recall the window decomposition $r_{f,m}^{\mathrm{PreTOM}} = \alpha_f^{\mathrm{PreTOM}} + d_f\, g_m + \varepsilon_{f,m}^{\mathrm{PreTOM}}$ and $r_{f,m}^{\mathrm{Rest}} = \alpha_f^{\mathrm{Rest}} + \varepsilon_{f,m}^{\mathrm{Rest}}$.

Summing the PreTOM decomposition over the trailing year ($m-12$ to $m-2$) gives the cash-window momentum signal

$$S_{f,m}^{\mathrm{PreTOM}} \;\equiv\; \sum_{j=2}^{12} r_{f,m-j}^{\mathrm{PreTOM}} \;=\; d_f\,\Lambda_m^- + \text{residual}, \qquad \Lambda_m^- \equiv \sum_{j=2}^{12} g_{m-j},$$

while the future cash-window return contains $d_f\, g_m$. Ignoring residual covariance, the month-m cross-sectional moments of a Fama–MacBeth regression of future on trailing cash-window returns are

$$\mathrm{Cov}_f\big(S_{f,m}^{\mathrm{PreTOM}},\, r_{f,m}^{\mathrm{PreTOM}}\big) \approx \Lambda_m^-\, g_m\, \mathrm{Var}_f(d_f), \qquad \mathrm{Var}_f\big(S_{f,m}^{\mathrm{PreTOM}}\big) \approx \big(\Lambda_m^-\big)^2\, \mathrm{Var}_f(d_f),$$

so the month-m slope is approximately g_m/Λ_m^- and the exposure dispersion $\mathrm{Var}_f(d_f)$ cancels. The $(\Lambda_m^-)^2$-weighted average that the pooled factor–month regression estimates has population slope

$$\gamma_{P\to P} \;\approx\; \frac{E[g_m\,\Lambda_m^-]}{E\big[(\Lambda_m^-)^2\big]}. \tag{14}$$

The denominator is a mean of squares and hence positive, so the sign is set by the numerator,

$$E\big[g_m\,\Lambda_m^-\big] \;=\; 11\,\bar g^2 \;+\; \sum_{j=2}^{12} \mathrm{Cov}(g_m,\, g_{m-j})\,.$$

The calendar component therefore arises from the recurring positive mean concession $\bar g > 0$ even without serial dependence, and persistence in g_m only strengthens it. Because the priced shock is absent from Rest returns, the same trailing signal does not predict future Rest returns, $\mathrm{Cov}_f(S_{f,m}^{\mathrm{PreTOM}},\, r_{f,m}^{\mathrm{Rest}}) \approx 0$, so $\gamma_{P\to P} > 0$ while $\gamma_{P\to R} \approx 0$.

Table IA.31: Theme-Collapsed Cross-Sectional Regression

Window	Slope	Bootstrap t	R^2	N
PreTOM	4.82	+3.99	0.86	14
Rest	0.27	+0.33	0.018	14

Notes: Cross-sectional regression of average daily long–short returns on dispensability exposure, equation (5), with the cross-section collapsed to the 14 JKP themes. Each observation is a theme: the equal-weighted mean, across the theme's factors, of the intended-signed average daily window return (bps/day) and of the intended-signed dispensability exposure d_f. The 52-week-high factor is excluded before collapsing, leaving 152 factors. Characteristics within a theme are built from overlapping inputs, so the factor-level regression could in principle overweight redundant variation; collapsing to themes removes this double-counting. Inference is a month-block bootstrap (B = 5,000 draws, seed 20260810): calendar months are resampled with replacement, per-factor window means and monthly dispensability exposures are recomputed inside each draw before collapsing to themes, and t is the point slope divided by the bootstrap standard deviation, matching the factor-level convention.

Table IA.32: Factor-Momentum Signal Decomposition: Which Past Window Predicts Which Future Window?

	Future PreTOM (bps/day)		Future Rest (bps/day)	
Trailing signal	Slope	t	Slope	t
trail 12–1, full month	+16.602	+4.42	+8.273	+3.19
trail 12–1, PreTOM only	+34.930	+4.70	+11.709	+2.02
trail 12–1, Rest only	+12.809	+2.78	+9.356	+3.13
d_f (dispensability exposure)	+5.931	+4.37	+0.860	+0.86

Notes: Fama–MacBeth cross-sectional regressions of next-month factor returns on each trailing signal, computed monthly across the JKP universe ($\sim$ 144 factors per month), with time-series Newey–West t-statistics at 12 monthly lags. Trailing signals are formed over months $m-12$ through $m-2$ (skip-one). Future targets are the next-month avg daily long–short return in the named within-month window, in bps/day. The d_f signal is the time-invariant per-factor dispensability loading; its cross-sectional ranking is identical to the rolling d_f-predicted past PreTOM premium.

IA.18 SYY Arbitrage-Asymmetry Robustness

This section asks whether the dispensability pattern is instead a manifestation of the Stambaugh, Yu, and Yuan (2015) (hereafter SYY) mispricing-arbitrage channel. The tests separate the two mechanisms. The SYY channel predicts stronger mispricing effects when overpricing interacts with limits to arbitrage, proxied by idiosyncratic volatility and sentiment. The liquidity-demand channel predicts a calendar-bound price concession that loads directly on dispensability.

We first reproduce SYY's headline result on their own data. Table IA.33 replicates their Table II, the value-weighted Fama–French three-factor alphas from a 5×5 independent sort on their individual-stock mispricing measure and idiosyncratic volatility. The idiosyncratic-volatility puzzle and its arbitrage asymmetry both appear. The high-minus-low volatility spread is -1.42% per month ($t = -7.1$) among the most overpriced stocks, $+0.25\%$ ($t = +1.3$) among the most underpriced, and -0.68% ($t = -5.0$) across all stocks, close to their reported -1.50, $+0.41$, and -0.78.

Table IA.33: Replication of Stambaugh, Yu, and Yuan (2015) Table II: idiosyncratic-volatility effects among underpriced versus overpriced stocks

Mispricing group	Highest IVOL	Lowest IVOL	High−Low
Most Overpriced	−1.74	−0.32	−1.42 (−7.1)
Next 20%	−0.65	−0.11	−0.54 (−2.7)
Next 20%	−0.15	−0.03	−0.12 (−0.6)
Next 20%	−0.11	+0.13	−0.25 (−1.3)
Most Underpriced	+0.38	+0.13	+0.25 (+1.3)
All stocks			−0.68 (−5.0)

Notes. Value-weighted Fama–French three-factor benchmark-adjusted returns (% per month) for the 5×5 independent sort of stocks on the Stambaugh, Yu, and Yuan (2015) mispricing measure and idiosyncratic volatility (the standard deviation of Fama–French three-factor daily residuals). We use their own individual-stock mispricing scores. Sample August 1965–January 2011; t-statistics in parentheses. Stambaugh, Yu, and Yuan (2015) report a High−Low IVOL spread of -1.50 ($t = -7.4$) for the most overpriced stocks, $+0.41$ ($t = +2.2$) for the most underpriced, and -0.78 ($t = -5.5$) for all stocks.

We then ask when in the month this relation operates and which stock trait organizes it. Table IA.34 reports Fama–MacBeth cross-sectional regressions of raw daily returns, estimated separately over the PreTOM window and the rest of the month. Idiosyncratic volatility predicts -1.4 basis points per day during PreTOM ($t = -2.6$) but is flat over the rest of the month (-0.1, $t = -0.2$). The volatility–return relation is thus a pre-turn-of-month phenomenon. Dispensability alone predicts -2.8 basis points per day ($t = -4.6$), and with both variables included idiosyncratic volatility falls to -0.4 ($t = -0.9$) while dispensability

retains -1.5 ($t = -3.1$). In the PreTOM cross-section, liquidation candidacy, not volatility, organizes returns, and volatility's own predictive power runs through it.[4] The mispricing-conditional part of the puzzle that SYY emphasize operates at a similar daily rate throughout the month, so we do not claim the entire puzzle is calendar-bound; the narrower statement is that the pre-turn-of-month discount on high-volatility stocks is a discount on dispensable stocks.

Table IA.34: Idiosyncratic volatility versus dispensability in the PreTOM cross-section

	PreTOM (bps/day per SD)	Rest (bps/day per SD)
Idiosyncratic volatility (alone)	-1.4 (-2.6)	-0.1 (-0.2)
Dispensability (alone)	-2.8 (-4.6)	$+0.3$ ($+0.8$)
Both, with mispricing and interactions		
Idiosyncratic volatility	-0.4 (-0.9)	$+0.0$ ($+0.2$)
Dispensability	-1.5 (-3.1)	$+1.5$ ($+3.8$)
IVOL $\times$ mispricing	-0.5 (-3.1)	-0.8 (-7.1)

Notes. Fama–MacBeth monthly cross-sectional regressions of each stock's average daily return (basis points per day) on cross-sectionally standardized idiosyncratic volatility, dispensability $d = -z(\text{prc}/52\text{H})$, the Stambaugh, Yu, and Yuan (2015) mispricing score, and interactions, estimated separately over the PreTOM window ($t-9$ to $t-4$) and the rest of the month. Coefficients are basis points per day per one-standard-deviation change; Newey–West(6) t-statistics in parentheses. Sample August 1965–January 2011. Returns are raw, not benchmark-adjusted, because the pricing factors themselves carry pre-turn-of-month components (see text).

Three further tests sharpen the discrimination: a stock-day triple interaction of dispensability with IVOL and the PreTOM dummy; a time-series regression of $\text{LDS}_m^{\text{PreTOM}}$ on lagged Baker–Wurgler sentiment; and a piecewise estimate of the PreTOM-versus-Rest return gap across dispensability percentile bins.

The triple interaction $D \times \text{IVOL} \times \text{PreTOM}$ captures the SYY prediction that dispensability and IVOL should amplify each other inside the cash-demand window. The PreTOM $\times D$ coefficient is the dominant loading ($t \approx -5$); the PreTOM $\times$IVOL coefficient is null; the triple interaction is marginal. The dispensability effect operates through D directly rather than through an interaction with IVOL.

SYY use the Baker-Wurgler sentiment index to identify time variation in mispricing. We replicate the same time-series logic with $\text{LDS}_m^{\text{PreTOM}}$ as the dependent variable. The sign

[4] We report raw cross-sectional slopes rather than benchmark-adjusted alphas. Because factor returns themselves carry pre-turn-of-month components, the size and value factors used for benchmark adjustment are no exception, so adjusting a PreTOM return by contemporaneous factor returns would absorb the effect under study. The question is in any case cross-sectional, namely which characteristic organizes PreTOM returns, and is answered directly by regressing returns on the competing characteristics.

Table IA.35: SYY arbitrage-asymmetry triple-interaction test (stock-day panel)

	Coef.	t-stat
Intercept	$+9.589^{***}$	$(+8.92)$
D (dispensability z-score)	$+1.019^{**}$	$(+2.49)$
IVOL	$+1.614^{***}$	$(+4.55)$
$D \times$ IVOL	$+3.142^{***}$	$(+12.23)$
PreTOM dummy	-10.669^{***}	(-5.26)
PreTOM $\times D$	-3.327^{***}	(-4.71)
PreTOM $\times$IVOL	$+0.946$	$(+1.36)$
PreTOM $\times D \times$ IVOL	-0.873^{*}	(-1.88)
R^2	0.0003	
N	55,343,910 stock-days	

Notes. Stock-day OLS panel regression of daily excess return (bps) on dispensability $D = -z(\text{prc}/52\text{H})$, $\text{IVOL}_{\text{FF3,21d}}$, their interaction, a PreTOM dummy, and the triple interaction $D \times \text{IVOL} \times \text{PreTOM}$. D and IVOL are cross-sectionally z-scored per month-end. Standard errors clustered by date. Sample: 1963–2025, CRSP $\times$ JKP characteristic universe. The PreTOM $\times D$ coefficient is the canonical dispensability concession; the triple interaction tests whether IVOL amplifies the dispensability effect inside PreTOM, the SYY arbitrage-asymmetry prediction. *,**,*** denote 10%, 5%, 1% significance.

is right but the slope is statistically indistinguishable from zero; the comparison regression on $\text{LDS}_m^{\text{Rest}}$ is comparable. Sentiment-driven mispricing is not the time-series driver of our calendar-bound shock.

Table IA.37 reports average daily returns by 20 dispensability percentile bins and trading-day window. The PreTOM-versus-Rest gap slopes monotonically negative across bins (slope -0.52 bps per bin, $t = -9.4$). The signature is the opposite of SYY's predicted mispricing-correction curve, which would be flat in the middle and steepen at both tails.

Table IA.36: SYY sentiment test: does Baker-Wurgler SENT_{t-1} predict $\text{LDS}_m^{\text{PreTOM}}$?

Dep. var.	Regressor	$\hat{\beta}$	t-stat (NW 12)	R^2
$\text{LDS}^{\text{PreTOM}}$	SENT_lag	$+3.646^{*}$	$(+1.67)$	0.0029
LDS^{Rest}	SENT_lag	$+4.441^{***}$	$(+2.71)$	0.0090
$\text{LDS}^{\text{PreTOM}}$	SENT_ORTH_lag	$+3.342$	$(+1.55)$	0.0024
LDS^{Rest}	SENT_ORTH_lag	$+4.070^{**}$	$(+2.48)$	0.0075

Notes. Newey–West (12-lag) time-series regression of $\text{LDS}_m^{\text{PreTOM}}$ (and the Rest placebo) on lagged Baker-Wurgler sentiment. `SENT_lag` is the raw BW index lagged one month; `SENT_ORTH_lag` is the macro-orthogonalized version. SYY predicts a positive coefficient on lagged sentiment: high sentiment → more overpricing of dispensable stocks → larger PreTOM concession. The pattern is sign-correct but insignificant; the placebo on LDS^{Rest} is comparable or stronger, indicating that sentiment-driven mispricing is not the PreTOM-specific channel. *,**,*** denote 10%, 5%, 1% significance.

Table IA.37: SYY piecewise test: average daily return by dispensability percentile and trading-day window

D-bin	Rest (bps/day)	t (NW 20)	PreTOM (bps/day)	t (NW 20)	Gap = PreTOM − Rest
1	$+7.58$	$(+9.54)$	$+1.42$	$(+1.14)$	-6.16
2	$+8.15$	$(+10.78)$	$+1.20$	$(+0.99)$	-6.96
5	$+9.29$	$(+11.09)$	$+0.98$	$(+0.75)$	-8.31
10	$+9.55$	$(+10.43)$	$+0.03$	$(+0.02)$	-9.51
15	$+9.68$	$(+8.39)$	-1.51	(-0.86)	-11.18
18	$+12.44$	$(+9.08)$	-2.49	(-1.24)	-14.93
19	$+14.86$	$(+10.54)$	-1.60	(-0.74)	-16.46
20	$+26.66$	$(+14.59)$	$+6.58$	$(+2.56)$	-20.08
Slope of Gap on D-bin		-0.521 bps/bin	$(t = -9.40)$		

Notes. Stocks sorted into 20 percentile bins per month-end on $D = -z(\text{prc}/52\text{H})$; bin 1 contains the least-dispensable stocks (closest to 52-week high) and bin 20 the most dispensable. Cell entries are the time-series mean of the bin's daily equal-weighted CRSP return (bps/day) within PreTOM days vs. Rest days, with Newey-West (20 lag) standard errors. The Gap column reports PreTOM mean minus Rest mean per bin. Only representative bins shown. The full 20-bin curve is in output/syy_piecewise_curve.csv. The negative slope of Gap on D-bin documents that the PreTOM-specific concession strengthens monotonically with dispensability. The slope-of-gap t in the final row is a cross-bin OLS statistic (a monotonicity diagnostic), not a Newey–West time-series statistic.

IA.19 Low-Risk and Lottery Factors

The low-risk anomaly, summarized by the BAB factor of Frazzini and Pedersen (2014), has a well-known leverage-constraint interpretation, in which leverage-constrained investors bid up high-beta stocks and depress their returns relative to low-beta stocks. High-beta, high-idiosyncratic-volatility, and high-maximum-daily-return "lottery" stocks are also disproportionately far from their 52-week highs, and so are among the dispensable holdings investors sell first in the PreTOM window. The leverage-constraint and liquidity-demand channels therefore operate on an overlapping set of stocks.

This overlap makes the low-risk family a stringent test, because market beta and dispensability are nearly collinear in precisely this corner of the cross-section. The two channels cannot be separated by a factor-level or risk-adjusted decomposition. Controlling for the market return removes much of the dispensability variation, so a CAPM-alpha test understates the liquidity component, while leaving beta uncontrolled lets it absorb the same variation. We therefore do *not* claim that the liquidity mechanism explains the low-risk premium, nor that the premium is independent of it; cross-sectionally, the leverage-constraint and liquidity-demand channels are not separately identified, as the relevant stocks are both high-beta and dispensable.

What can be identified is whether a dispensability concession operates *within* the low-risk leg, holding factor-leg membership fixed and netting residual risk characteristics. We pool the short (high-risk) legs of the twenty JKP low-risk factors and regress each stock's PreTOM-minus-Rest market-adjusted return on its within-leg dispensability (the within-month z-score of distance below the 52-week high, $\delta_{i,m}$), absorbing factor-by-month fixed effects with standard errors clustered by month. The concession is positive for all twenty factors individually; the twelve whose per-factor estimates are individually insignificant are jointly significant when pooled; and the estimate *strengthens* when we add continuous controls for beta, idiosyncratic volatility, prior return, maximum daily return, and size ($\lambda = 6.9$ bps/day, $t = 6.2$). It is unchanged by six- and twelve-month momentum controls and attenuates but remains significant net of one-month reversal ($\lambda = 4.0$ bps/day, $t = 3.7$). Holding low-risk leg membership fixed and controlling for residual beta and volatility, the more dispensable stocks pay an additional PreTOM concession.

This identifies a calendar-specific liquidity-demand component inside the high-risk leg itself, a six-day, settlement-timed concession that the leverage-constraint mechanism does not naturally predict, since leverage constraints carry no month-end clock. It does not establish that the low-risk premium *is* liquidity demand. The leverage-constraint channel may still explain why high-beta and lottery stocks earn low average returns, and our within-leg

evidence identifies a distinct component rather than a replacement. We note that the corresponding order-flow test (whether dispensable stocks within the low-risk leg draw heavier seller-initiated volume in PreTOM) is not statistically significant in the intraday-trade sample (2003–2022), so the within-leg identification rests on returns and their calendar signature rather than on directly observed selling.

IA.20 Factor-Construction Robustness

Our baseline analysis uses decile-sorted portfolios that we construct from the 153 firm characteristics of Jensen, Kelly, and Pedersen (2023), with NYSE breakpoints, value weights, and a uniform $D_{10} - D_1$ sign convention applied to every characteristic. We adopt this construction because it delivers a transparent short-minus-long mapping from stock-level dispensability to factor returns, and because it matches the decile-style libraries of Chen and Zimmermann (2022) and Hou, Xue, and Zhang (2020) against which we cross-validate. The cross-sectional result is unchanged under Jensen, Kelly, and Pedersen's own published portfolios and two independent external libraries, and our from-primitives construction reproduces the official JKP portfolios closely.

AB.0. Distribution of Dispensability Exposure

Figure IA.4 plots the cross-sectional distribution of the factor-level raw dispensability gap $\widetilde{d}_f$ across the Jensen, Kelly, and Pedersen (2023) factors (the unsigned $D1$-minus-$D10$ analog of Equation (7)), excluding the 52-week-high characteristic from which it is built. The gap is centered near zero (mean +0.15, cross-sectional standard deviation 0.46), with 69% of factors short more-dispensable stocks than they are long ($\widetilde{d}_f > 0$). The dispersion is wide relative to the mean: momentum and the profitability factors sit far on the positive (plumbing) side, while value, distress, and idiosyncratic-volatility factors load negatively. This is the cross-sectional variation that the PreTOM regressions exploit.

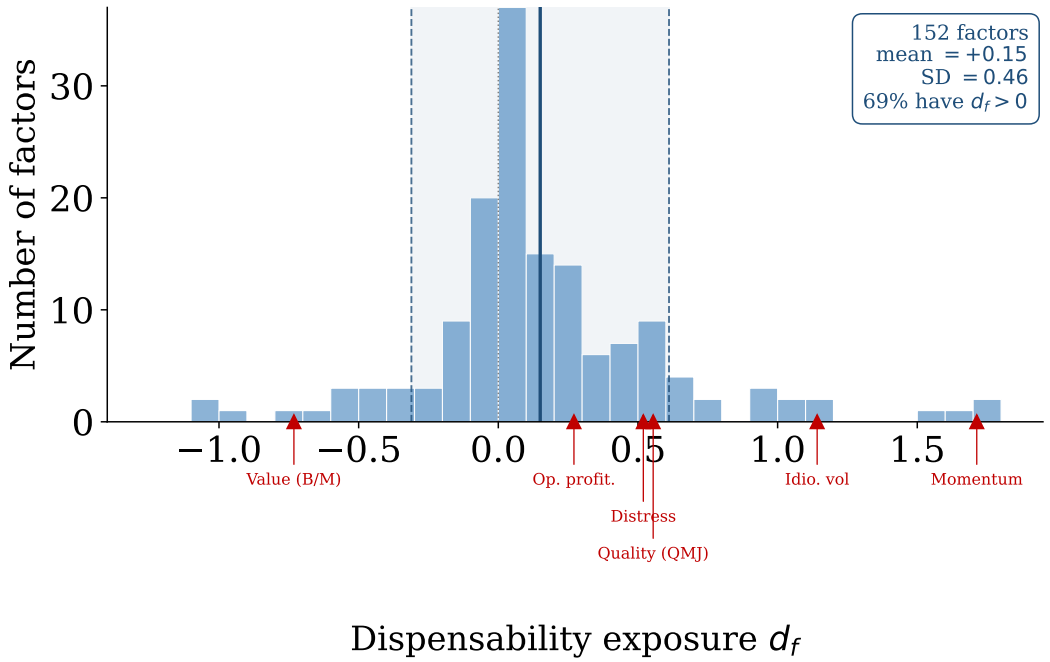

Figure IA.4: Cross-Sectional Distribution of the Raw Dispensability Gap $\widetilde{d}_f$. Histogram of the factor-level raw dispensability gap $\widetilde{d}_f = \frac{1}{M}\sum_m(\bar{\delta}_{1,f,m} - \bar{\delta}_{10,f,m})$, the unsigned $D1$-minus-$D10$ analog of Equation (7) and the value-weighted $D1$-minus-$D10$ average of the stock-level 52-week-high dispensability score, across 152 Jensen, Kelly, and Pedersen (2023) factors. The construction proxy (`prc_highprc_252d`) is excluded because its gap is mechanically its own $D1$-$D10$ spread. The solid line marks the mean (+0.15); the dashed lines and shaded band mark ±1 cross-sectional standard deviation (0.46). The horizontal axis is the short-minus-long 52-week-high score. Triangles mark representative factors. A positive $\widetilde{d}_f$ means the factor is short the more-dispensable side, so month-end selling lifts its return; 69% of factors have $\widetilde{d}_f > 0$.

AB.1. Alternative Factor Constructions

Table IA.38 re-estimates the main cross-sectional regression of each factor's average PreTOM (and Rest) daily return on its dispensability exposure, under four distinct factor-return constructions. The first row is our baseline author-built decile spread, with the holdings-based exposure d_f. The second keeps the holdings-based d_f but replaces the returns with Jensen, Kelly, and Pedersen's *official* low/high tercile portfolios (their capped value-weighted `ret_vw_cap` series, high tercile minus low tercile), so that the factor returns are entirely theirs. The third and fourth rows use the decile libraries of Chen and Zimmermann (2022) and Hou, Xue, and Zhang (2020), built by independent research teams from their own code and breakpoints, taking their returns as provided and estimating exposure from return time series (b^{52H}) rather than from our portfolio assignments.

Table IA.38: Factor Construction Does Not Drive the Result

Construction	Returns used	Exposure used	PreTOM R^2	Rest R^2
Baseline	Author $D_{10} - D_1$ (JKP chars.)	Holdings d_f	0.724	0.010
Official JKP	Official low/high terciles	Holdings d_f	0.246	0.067
Chen–Zimmermann	As provided	Return-based b^{52H}	0.628	0.019
Hou–Xue–Zhang	As provided	Return-based b^{52H}	0.714	0.103

Notes: Each row reports the cross-sectional R^2 from regressing factors' average daily long-short return on their dispensability exposure, separately for the PreTOM window $[\tau-9, \tau-4]$ and the Rest of the month, under a different factor-return construction. The first two rows use the holdings-based exposure d_f; the external libraries use the return-based analog b^{52H} (Section 3.7). The 52-week-high characteristic that defines d_f is excluded, leaving 152 JKP characteristics; the external libraries use their full factor sets. Official JKP portfolios are the published capped value-weighted low/high terciles from Jensen, Kelly, and Pedersen (2023). The PreTOM relation is stronger than the Rest relation in every construction, ranging from roughly 3.7× on the official JKP terciles (0.246 vs. 0.067) and about 7× on Hou–Xue–Zhang to more than an order of magnitude on the author-built deciles and Chen–Zimmermann.

The PreTOM-versus-Rest asymmetry remains significant in all four constructions. It is weaker on the official JKP terciles than on our deciles only because terciles are mechanically less extreme than deciles (the top and bottom thirds of the cross-section span a narrower range of dispensability than the top and bottom deciles), yet even there the PreTOM relation is roughly 3.7 times stronger than the Rest relation. The two independent libraries, built without any of our code, reproduce the baseline closely. The PreTOM asymmetry is present under author-built JKP deciles, official JKP terciles, and two independently constructed anomaly libraries.

AB.2. We Reproduce the Official JKP Portfolios

We also verify that our from-primitives construction reproduces Jensen, Kelly, and Pedersen's published portfolios when we follow their recipe exactly. Jensen, Kelly, and Pedersen (2023) form breakpoints from the equal-count thirds of all *non-micro* stocks in a country (those above the NYSE 20th percentile of market equity), then distribute the micro-cap stocks into those same three groups; each tercile return is a capped value weight, weighting by market equity winsorized at the NYSE 80th percentile. Rebuilding the terciles from the underlying characteristics under this exact recipe and correlating our daily high-minus-low return with their official `ret_vw_cap` spread, characteristic by characteristic, yields the distribution in Table IA.39.

Table IA.39: Replication of Official JKP Portfolios from Primitives

Characteristic	N (days)	Corr. with official
Momentum (12–1)	15,606	0.999
Book-to-market	15,606	0.999
Asset growth	15,606	0.999
Operating profitability	15,606	0.998
Gross profitability	15,606	0.998
Market beta (60m)	15,606	0.999
Quality-minus-junk	15,606	0.998
Median, all 153 characteristics	—	*0.999*
10th percentile	—	*0.993*
Share with corr. > 0.95	—	*95%*
Share with corr. > 0.90	—	*97%*

Notes: Daily time-series correlation between our from-primitives high-minus-low tercile return and Jensen, Kelly, and Pedersen (2023)'s official capped value-weighted (`ret_vw_cap`) spread, by characteristic, over 1963–2025. Our terciles are built using JKP's exact recipe: equal-count thirds of non-micro stocks for breakpoints, with micro caps distributed into the same groups, and capped value weights (market equity winsorized at the NYSE 80th percentile). The top panel lists canonical factors; the bottom panel reports the distribution across all 153 characteristics.

The median characteristic correlates with its official JKP counterpart at 0.999, and 95% of characteristics exceed 0.95. Every economically central factor—momentum, value, asset growth, profitability, beta, and quality—reproduces above 0.97. The handful of low-correlation cases are the quarterly-data growth and earnings-change characteristics (`lti_gr1a`, `sti_gr1a`, `inv_gr1a`, `niq_at_chg1`, `niq_be_chg1`), which are not simple sorts but depend on Jensen, Kelly, and Pedersen's annualization and standardized-change helper functions; excluding them leaves the estimate unchanged. The residual gap on the remaining factors

is irreducible independent-rebuild noise from delisting-return handling, the exact per-date stock universe, and data vintage, not a difference in the sorting recipe.

IA.21 Factor-Level Dispensability Exposure

The factor-level classification used in the paper is the exposure taxonomy in Table 5 of the main paper, which ranks the 152 tested JKP factors by their dispensability exposure d_f and flags the low-exposure factors (18 of 152) whose returns are essentially orthogonal to the PreTOM concession. Per-factor exposures, PreTOM and Rest daily long–short returns, and their t-statistics for all 152 factors are provided in the replication package.

IA.22 Has the Observable Proxy for Dispensability Changed Over Time?

Table IA.40 reports the cross-factor PreTOM R^2 separately for 1963–1993 and 1994–2025 for the leading candidate characteristics. Price-history and trading-friction proxies weaken while quality, safety, and fundamental-distress proxies strengthen: distance below the 52-week high falls from 0.68 to 0.45, whereas the Stambaugh–Yuan performance composite rises from 0.11 to 0.60, quality-minus-junk from 0.17 to 0.57, the F-score from 0.15 to 0.50, and QMJ safety from 0.23 to 0.49; the bid-ask spread (0.41 $\rightarrow$ 0.21) and CAPM idiosyncratic volatility (0.45 $\rightarrow$ 0.18) decline. Figure IA.5 plots the same R^2s on rolling 15- and 20-year windows. The crossover between the 52-week-high proxy and the quality and mispricing proxies is gradual and centered on the 2000s, while the bid-ask proxy alone breaks around decimalization in 2001.

Table IA.40: Cross-Factor PreTOM R^2 by Subperiod: Rotation of the Observable Proxy

Characteristic	1963–1993	1994–2025
Distance from 52-week high	0.68	0.45
Mispricing (performance)	0.11	0.60
Quality-minus-junk (QMJ)	0.17	0.57
F-score	0.15	0.50
QMJ safety	0.23	0.49
Bid-ask spread	0.41	0.21
CAPM idiosyncratic volatility	0.45	0.18

Table IA.41 runs the decisive joint regression $\bar{R}_f^{\text{PreTOM}} = a + b_{52}\, d_f^{52\text{H}} + b_Q\, d_f^{Q} + u_f$ separately by half. Early, distance from the high absorbs the quality dimension: adding QMJ (or PERF, or the F-score) leaves the joint R^2 at 0.68 and the quality partial R^2 near zero ($t < 1.2$). Late, the relation becomes genuinely two-dimensional: the joint R^2 (0.68) exceeds both univariate fits, quality enters significantly (QMJ $t = +2.45$, partial $R^2 = 0.43$), and distance from the high retains a large partial R^2 (0.26) although its incremental t (+1.33) weakens under collinearity. This rejects both a clean proxy swap (the 52-week high does not vanish) and pure correlation (the joint fit rises materially above the best univariate).

The declining fit of the 52-week-high proxy does not indicate a weakening mechanism. The factor-level exposure is highly persistent (cross-half correlation 0.958), early-period exposure predicts the later PreTOM cross-section at roughly $R^2 = 0.51$ with negligible Rest fit, and the modern-sample tests that use the 52-week-high proxy all remain present: the 2003–2025 PreTOM–Rest asymmetry, TAQ seller-initiated order flow (2003–2022), the 2024 settlement shift, the ETF cross-section, and the within-leg spreads. Liquidation candidacy

Table IA.41: Joint Cross-Factor PreTOM Regression on 52-Week-High and Quality Exposure

Period	Pair	52H alone	Quality alone	Joint R^2
Early	52H + QMJ	0.68	0.17	0.68
Late	52H + QMJ	0.45	0.57	0.68
Early	52H + PERF	0.67	0.11	0.68
Late	52H + PERF	0.44	0.60	0.67
Early	52H + F-score	0.68	0.16	0.69
Late	52H + F-score	0.45	0.50	0.60

Notes: Cross-factor regression of mean signed PreTOM factor returns on factor-level 52-week-high exposure and one quality exposure, leave-one-out over 152 factors, month-block bootstrap. In the late half the QMJ partial R^2 is 0.43 ($t = +2.45$) and the 52H partial R^2 is 0.26 ($t = +1.33$).

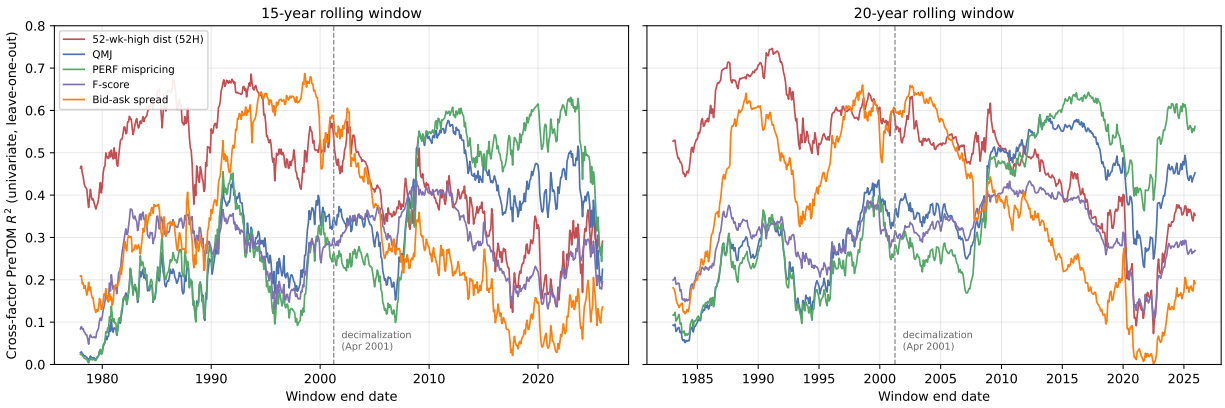

Figure IA.5: Rolling cross-factor PreTOM R^2 for five candidate characteristics (15- and 20-year windows). The crossover between distance below the 52-week high and the quality/mispricing proxies is gradual and centered on the 2000s; the bid-ask proxy alone breaks around decimalization (2001).

is the economic primitive; distance below the 52-week high is its strongest parsimonious full-sample proxy, and fundamental quality and safety have become increasingly informative in recent decades.

One might ask whether the headline should therefore use a composite proxy that also loads on quality. Combining the two raises the modern-period fit only modestly and changes no economic conclusion. Adding QMJ to distance from the high lifts the cross-factor PreTOM R^2 from 0.45 (52-week high alone) to 0.68 over 1994–2025, but the increment over the best single proxy (QMJ, 0.57) is only 0.11, and it shrinks to 0.04 in the stricter 2003–2025 window, where the two dimensions are collinear and neither is individually significant (52-week-high partial $t = 0.74$, QMJ $t = 1.73$). An equal-weight composite of distance-from-high, QMJ safety, and gross profitability behaves the same way: it fits marginally better late (0.46 in 2003–2025) but *worse* over the full sample (0.59 versus 0.72), because fixed equal weighting dilutes the dominant proxy. Throughout, the slope stays positive—dispensable and junk stocks underperform in PreTOM—and the PreTOM-versus-Rest asymmetry is preserved or slightly sharper, with the composite's Rest R^2 at 0.01 or below. Combining proxies thus raises fit slightly in recent decades but adds no new economic content, so we retain the parsimonious single-characteristic proxy in the main text and treat the growing role of quality as a robustness result rather than a respecification.

Table IA.42: Robustness to Price and Size Screens

Universe	PreTOM R^2	Rest R^2	Mean LDS (bps/day)	d_f×LDS β (t)
Baseline	0.72	0.015	+8.9	0.50 (16.7)
Price $\geq$ \$5	0.67	0.007	+7.3	0.47 (14.7)
Above NYSE 20th size pctile	0.65	0.001	+6.5	0.46 (14.9)
Both restrictions	0.63	0.000	+6.6	0.45 (14.6)

Notes: Each screen is applied to the stock universe before sorting (governing both NYSE breakpoints and decile membership); all statistics are recomputed on the screened population. The headline PreTOM fit attenuates modestly while the Rest fit falls toward zero and the d_f×LDS panel loading stays large and highly significant.

IA.23 Composition of the Dispensable-Stock Universe

Table IA.43 profiles the stocks in the most-dispensable decile against the least-dispensable decile and the full universe, over the full sample and by era.

Table IA.43: What Are Dispensable Stocks? Composition of the Most-Dispensable Decile

Characteristic	Full sample 1963–2025			Most-dispensable decile	
	Most disp.	Least disp.	Universe	1963–1993	1994–2025
Piotroski F-score (median)	3.8	5.3	4.9	4.0	3.6
ROE, ni/be (median)	-0.21	0.11	0.08	0.02	-0.42
ROIC, ebit/bev (median)	-0.11	0.15	0.12	0.07	-0.28
Momentum rank (pctile, mean)	15	72	50	17	13
Share with negative 12m return (%)	82	8	42	78	85
Idiosyncratic vol., daily (median)	0.053	0.019	0.027	0.044	0.062
Market beta (median)	1.52	0.94	1.11	1.55	1.49
Size rank (pctile, mean)	24	62	50	26	22
Dividend yield (median)	0.001	0.013	0.009	0.001	0.000
Turnover, 126d (median)	0.007	0.003	0.003	0.002	0.012

Notes: Stocks are sorted monthly into deciles of the dispensability score $\delta_{i,m}$ (distance below the 52-week high, within-month standardized). Cells are time averages of monthly cross-sectional statistics (medians for ratio-type characteristics, means for percentile ranks and shares). The era columns show the most-dispensable decile by subperiod. Its fundamentals deteriorate in the recent era, and the deepening survives netting out the universe drift. The decile-minus-universe median-ROE gap widens from -8 to -48 percentage points, while the universe median moved only from 10.8% to 5.8%. Carbon-intensity composition is deferred (emissions coverage begins around 2005).

References

Chen, A. Y., and T. Zimmermann. 2022. Open source cross-sectional asset pricing. *Critical Finance Review* 11:207–264.

Coval, J., and E. Stafford. 2007. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics* 86:479–512.

Bai, J., and S. Ng. 2002. Determining the number of factors in approximate factor models. *Econometrica* 70:191–221.

Onatski, A. 2010. Determining the number of factors from empirical distribution of eigenvalues. *Review of Economics and Statistics* 92:1004–1016.

Ahn, S. C., and A. R. Horenstein. 2013. Eigenvalue ratio test for the number of factors. *Econometrica* 81:1203–1227.

Green, J., J. R. M. Hand, and X. F. Zhang. 2017. The characteristics that provide independent information about average U.S. monthly stock returns. *Review of Financial Studies* 30:4389–4436.

Feng, G., S. Giglio, and D. Xiu. 2020. Taming the factor zoo: A test of new factors. *Journal of Finance* 75:1327–1370.

Frazzini, A., and L. H. Pedersen. 2014. Betting against beta. *Journal of Financial Economics* 111:1–25.