

# Zwei-Stufen-XGBoost – Experiment-Report

Experiment-ID: hp\_long\_final\_yahoo

Dieses Dokument fasst die wichtigsten Parameter, Datenquellen und Metriken eines Zwei-Stufen-XGBoost-Experiments zusammen.

Stufe 1 (Signal): neutral vs. Bewegung ('move'). Stufe 2 (Richtung): down vs. up – nur an Bewegungstagen.

## Label-Parameter:

- horizon\_days: 4
- up\_threshold: 0.004
- down\_threshold: -0.004
- strict\_monotonic: False
- max\_adverse\_move\_pct: 0.01
- hit\_within\_horizon: True (True = Schwelle reicht, wenn sie irgendwo im Horizont erreicht wird)

## Datensatz & Splits:

- dataset\_path:  
  
data/processed/datasets/eurusd\_news\_training\_\_hp\_long\_final\_yahoo.csv
- test\_start: 2025-01-01
- train\_frac\_within\_pretest: 0.8
- feature\_mode: price\_only

## Entscheidungsgrenzen (Modelle):

- SIGNAL\_THRESHOLD (Stufe 1 – move vs. neutral): 0.5 (höher → höhere Precision, niedrigerer Recall).
- SIGNAL\_THRESHOLD\_TRADE (Stufe 1 – Trading): 0.425 (höher → weniger Trades, tendenziell höhere Qualität).
- DIRECTION\_THRESHOLD (Stufe 2 – down vs. up, für Metriken): 0.42000000000000004 (niedriger → mehr up, höher → weniger up).
- DIRECTION\_THRESHOLDS (Stufe 2 – Trading-Entscheidungen): down, wenn  $P(\text{up}) \leq 0.3$ , up, wenn  $P(\text{up}) \geq 0.6749999999999999$ .

# Legende & Begriffe (Kurzüberblick)

## Zielvariablen:

- label: 3-Klassen-Ziel auf Basis des 4-Tage-Lookaheads (neutral / up / down).
- signal: 0 = neutral, 1 = Bewegung (up oder down).
- direction: 0 = down, 1 = up; nur definiert, wenn signal == 1.

## Wichtige Metriken:

- precision: Anteil der vorhergesagten positiven Fälle, die wirklich positiv sind.
- recall: Anteil der tatsächlichen positiven Fälle, die erkannt wurden.
- f1: harmonischer Mittelwert aus precision und recall (Balance beider Größen).
- support: Anzahl der Beobachtungen in der jeweiligen Klasse.

## Feature-Abkürzungen (Auswahl, nicht vollständig – vollständige Liste siehe Seite 'Verwendete

### Features'):

- article\_count: Anzahl News-Artikel pro Tag.
- avg\_polarity / avg\_neg / avg\_neu / avg\_pos: durchschnittliche Sentiment-Werte.
- pos\_share / neg\_share: Anteil positiver bzw. negativer Sentiment-Komponente.
- intraday\_range\_pct: (High - Low) / Close – relative Tages-Spanne (Volatilität).
- upper\_shadow / lower\_shadow: obere/untere Dochte der Kerzen (High/Low vs. Körper).
- month / quarter: Kalendermonat und Quartal.

## Modell-Parameter (XGBoost)

### Signal-Modell (Stufe 1):

- objective: binary:logistic
- max\_depth: 3
- learning\_rate: 0.05
- n\_estimators: None
- subsample: 0.9
- colsample\_bytree: 0.9
- scale\_pos\_weight: 0.15201768933112217

### Richtungs-Modell (Stufe 2):

- objective: binary:logistic
- max\_depth: 3
- learning\_rate: 0.05
- n\_estimators: None
- subsample: 0.9
- colsample\_bytree: 0.9
- scale\_pos\_weight: 1.0

## Verwendete Features (FEATURE\_COLS)

#	feature_name	description
0	intraday_range_pct	$(\text{High} - \text{Low}) / \text{Close}$ – relative Tagesvolatilität.
1	upper_shadow	Oberer Kerzendocht: $\text{High} - \max(\text{Open}, \text{Close})$ .
2	lower_shadow	Unterer Kerzendocht: $\min(\text{Open}, \text{Close}) - \text{Low}$ .
3	price_close_ret_1d	Relativer Schlusskurs-Return gegenüber Vortag: $\text{Close}_t / \text{Close}_{\{t-1\}} - 1$ .
4	price_close_ret_5d	Schlusskurs-Return über 5 Tage: $\text{Close}_t / \text{Close}_{\{t-5\}} - 1$ .
5	price_range_pct_5d_std	Standardabweichung der intraday_range_pct über 5 Tage (Volatilität).
6	price_body_pct_5d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 5 Tage.
7	price_close_ret_30d	Schlusskurs-Return über 30 Tage: $\text{Close}_t / \text{Close}_{\{t-30\}} - 1$ .
8	price_range_pct_30d_std	Standardabweichung der intraday_range_pct über 30 Tage.
9	price_body_pct_30d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 30 Tage.
10	month	Kalendermonat (1–12).
11	quarter	Kalenderquartal (1–4).
12	cal_dow	Wochentag (0 = Montag, 6 = Sonntag).
13	cal_day_of_month	Kalendertag im Monat.
14	cal_is_monday	Flag: 1 wenn Montag, sonst 0.
15	cal_is_friday	Flag: 1 wenn Freitag, sonst 0.
16	cal_is_month_start	Flag: 1 wenn Monatsanfang, sonst 0.
17	cal_is_month_end	Flag: 1 wenn Monatsende, sonst 0.
18	hol_is_us_federal_holiday	Flag: 1 wenn US-Feiertag, sonst 0.
19	hol_is_day_before_us_federal_holiday	Flag: 1 wenn Tag vor US-Feiertag.
20	hol_is_day_after_us_federal_holiday	Flag: 1 wenn Tag nach US-Feiertag.

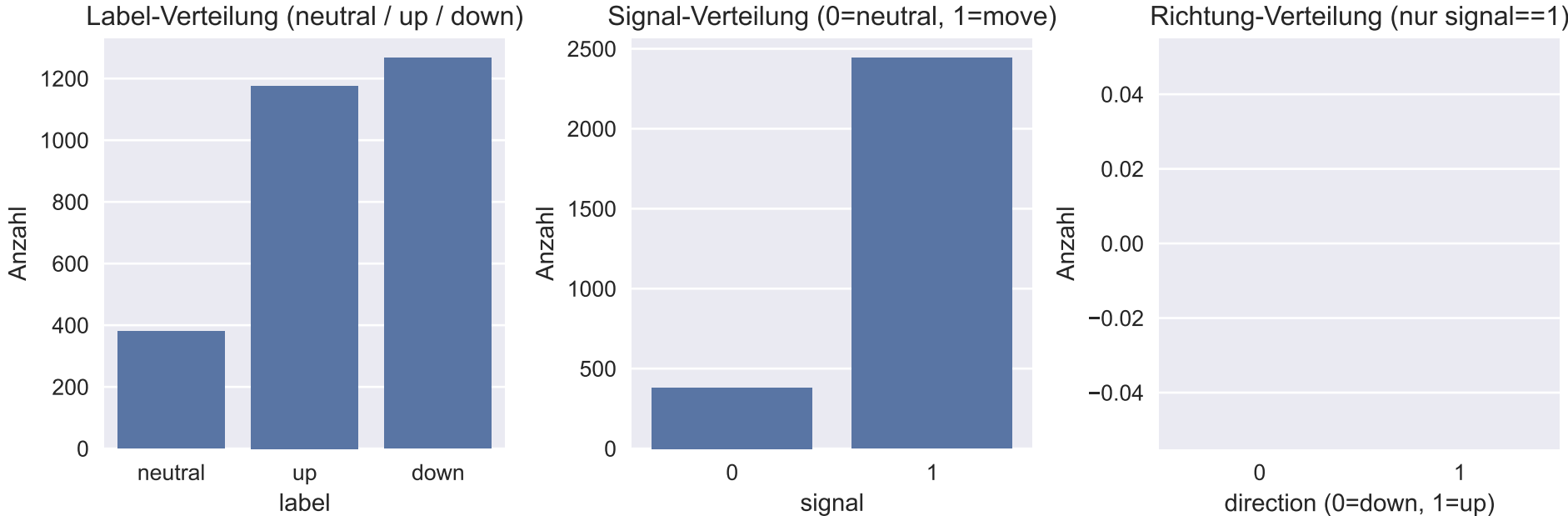
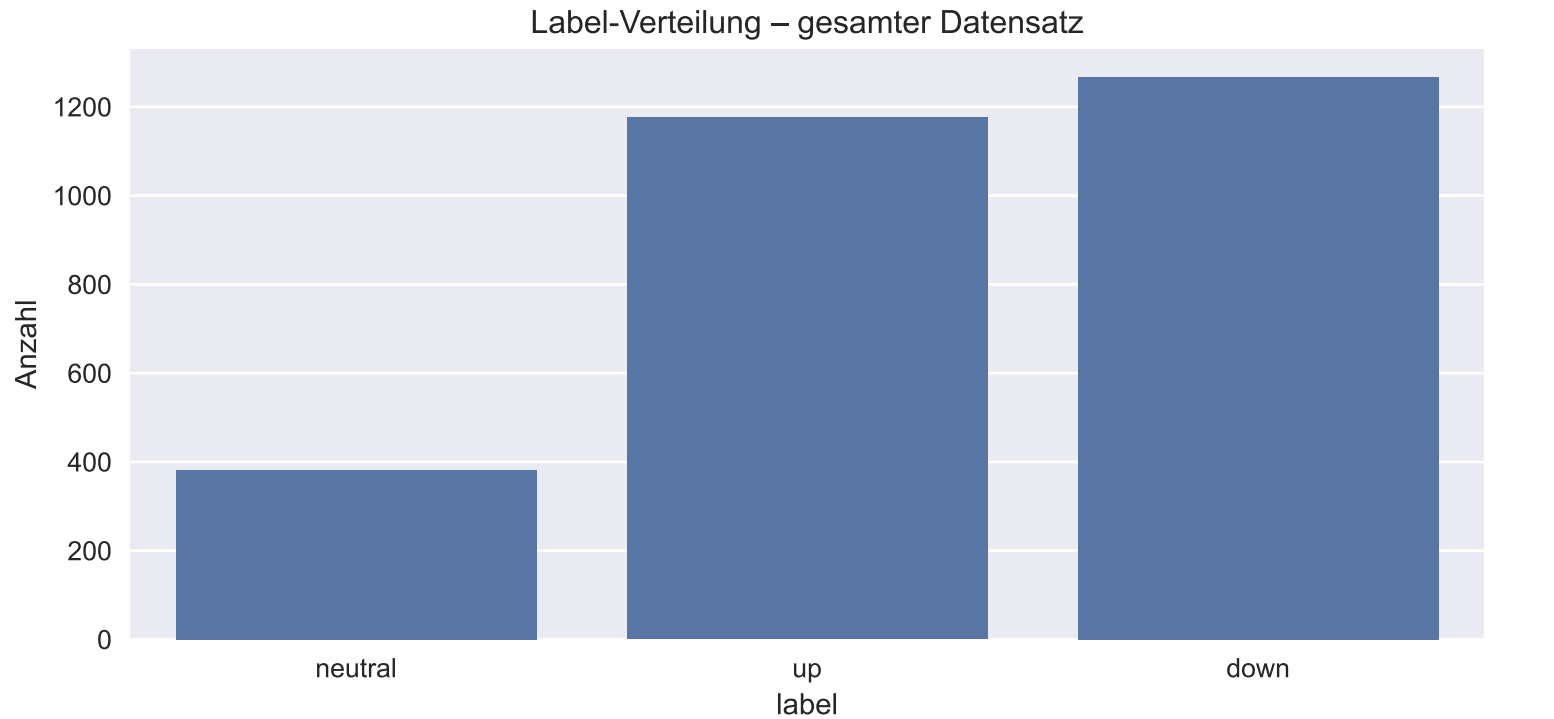
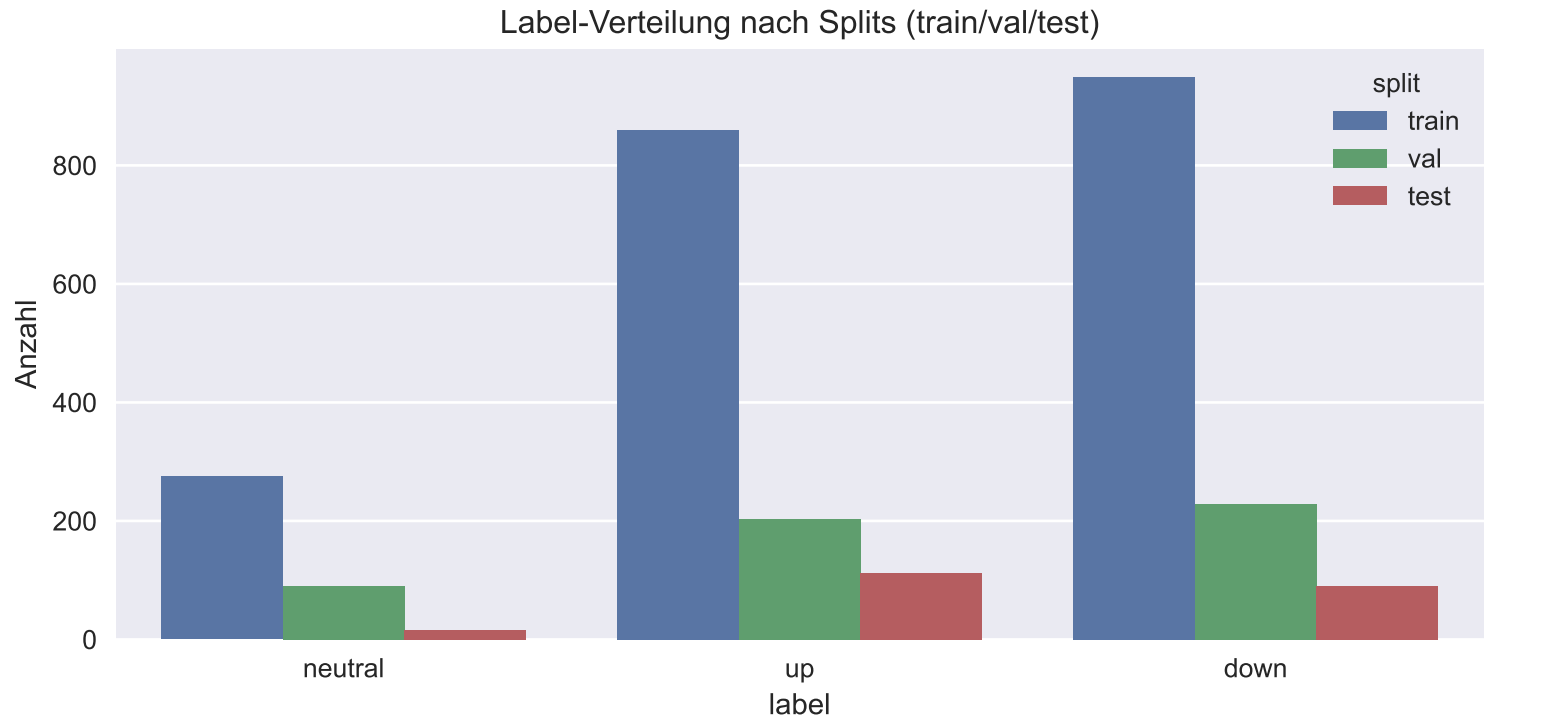


Abbildung: Klassenverteilungen für label, signal und direction im vollständigen Trainingsdatensatz.



label	count
neutral	381
up	1176
down	1267

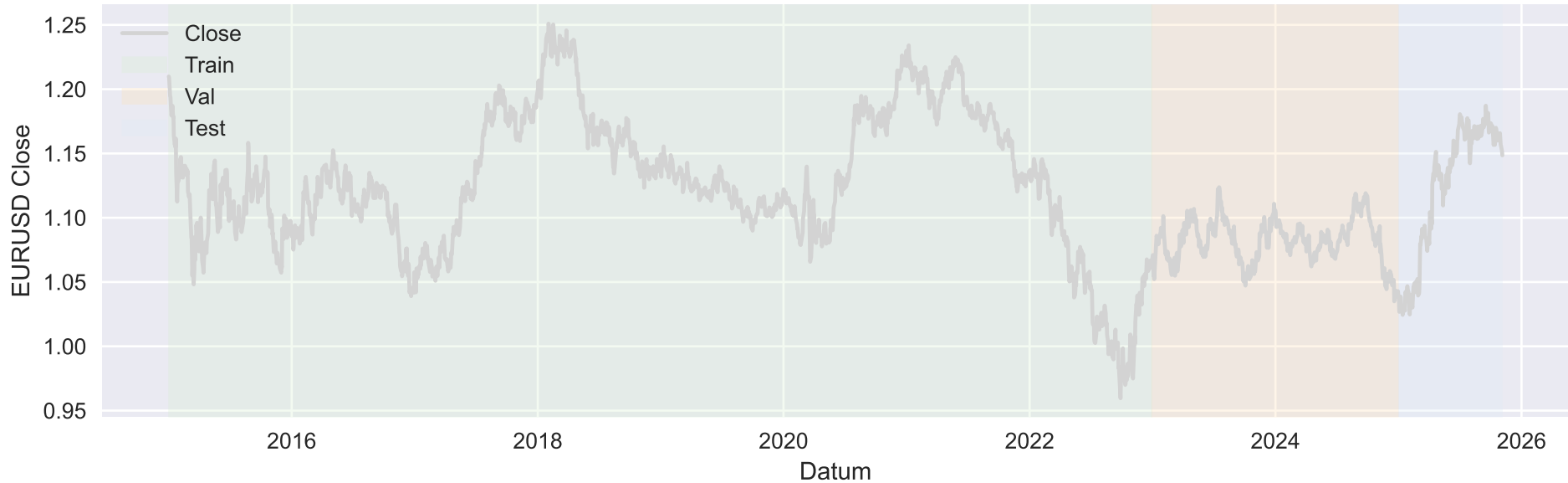
Abbildung/Tabelle: Verteilung der Zielvariable 'label' (neutral/up/down) im gesamten Datensatz.



split	neutral	up	down
train	275	860	949
val	90	204	228
test	16	112	90

Abbildung/Tabelle: Label-Verteilung getrennt nach Trainings-, Validierungs- und Test-Split.

EURUSD-Zeitreihe mit Train/Val/Test-Bereichen

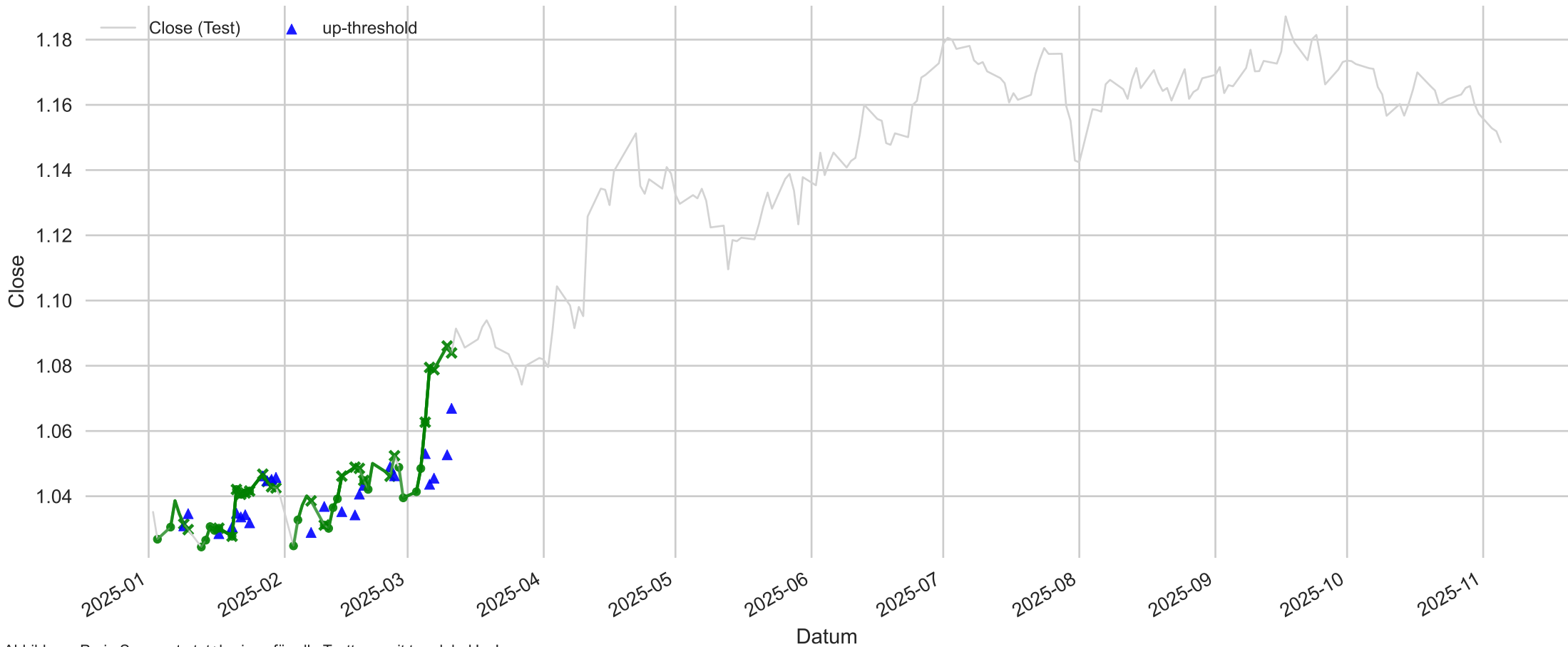


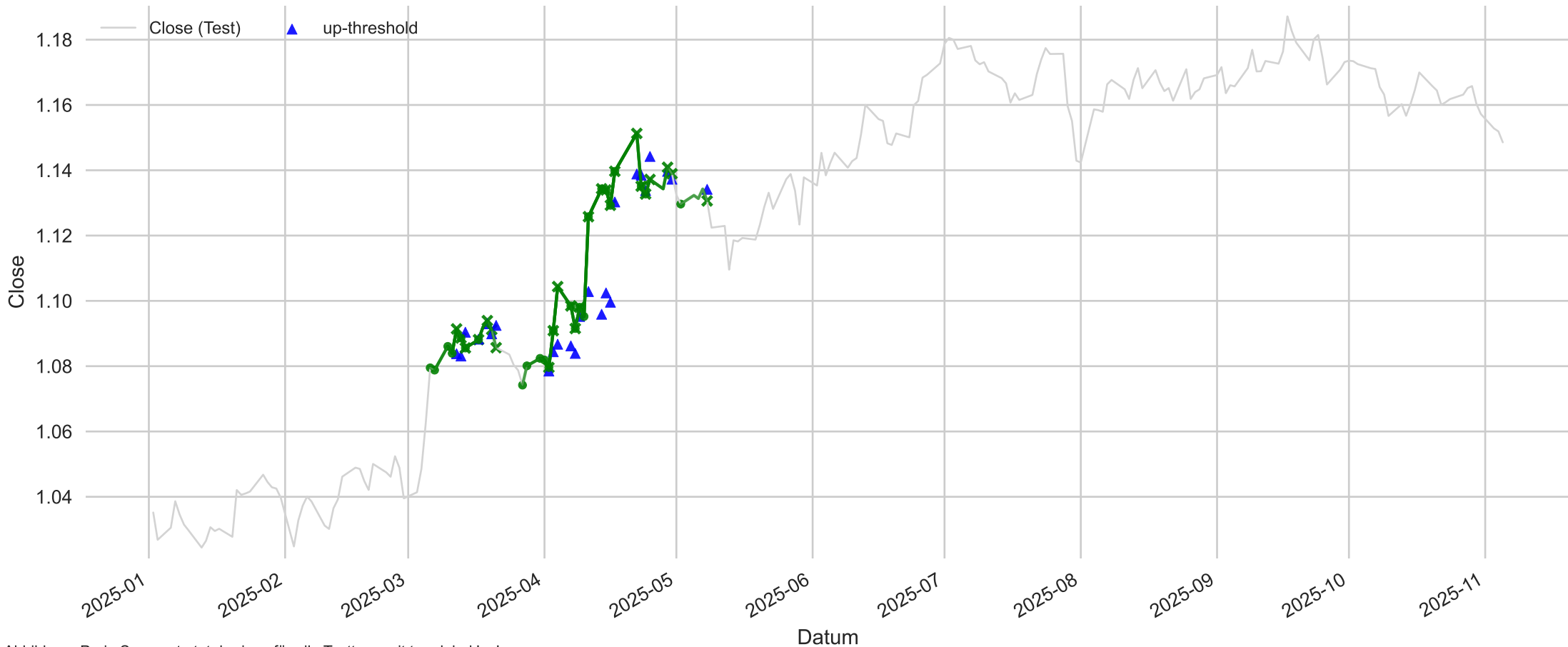


EURUSD-Zeitreihe mit hervorgehobenen up/down-Tagen (ab 2020)



Abbildung: EURUSD-Schlusskurs mit markierten up-/down-Tagen im betrachteten Zeitraum.





# EURUSD-Segmente mit label='up' (Test-Split) – Seite 3/5

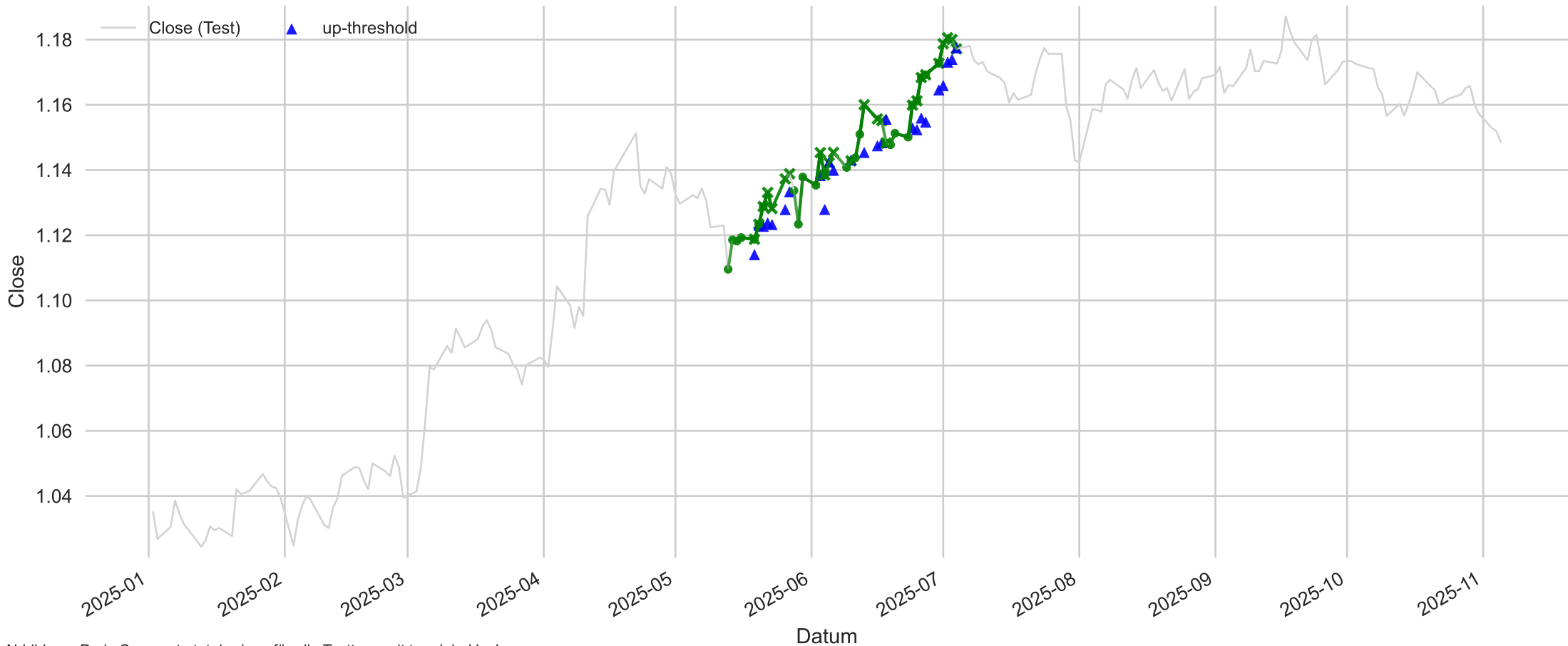
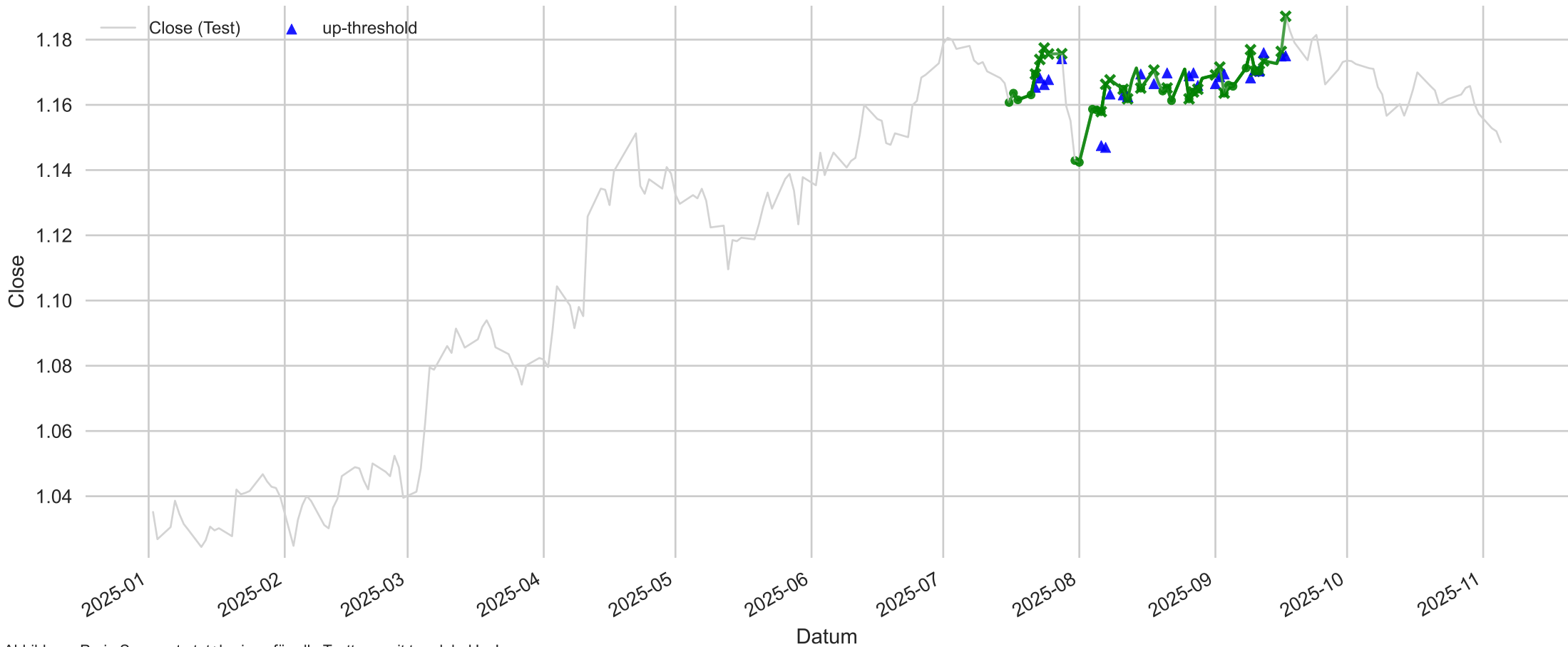


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'up'.



# EURUSD-Segmente mit label='up' (Test-Split) – Seite 5/5

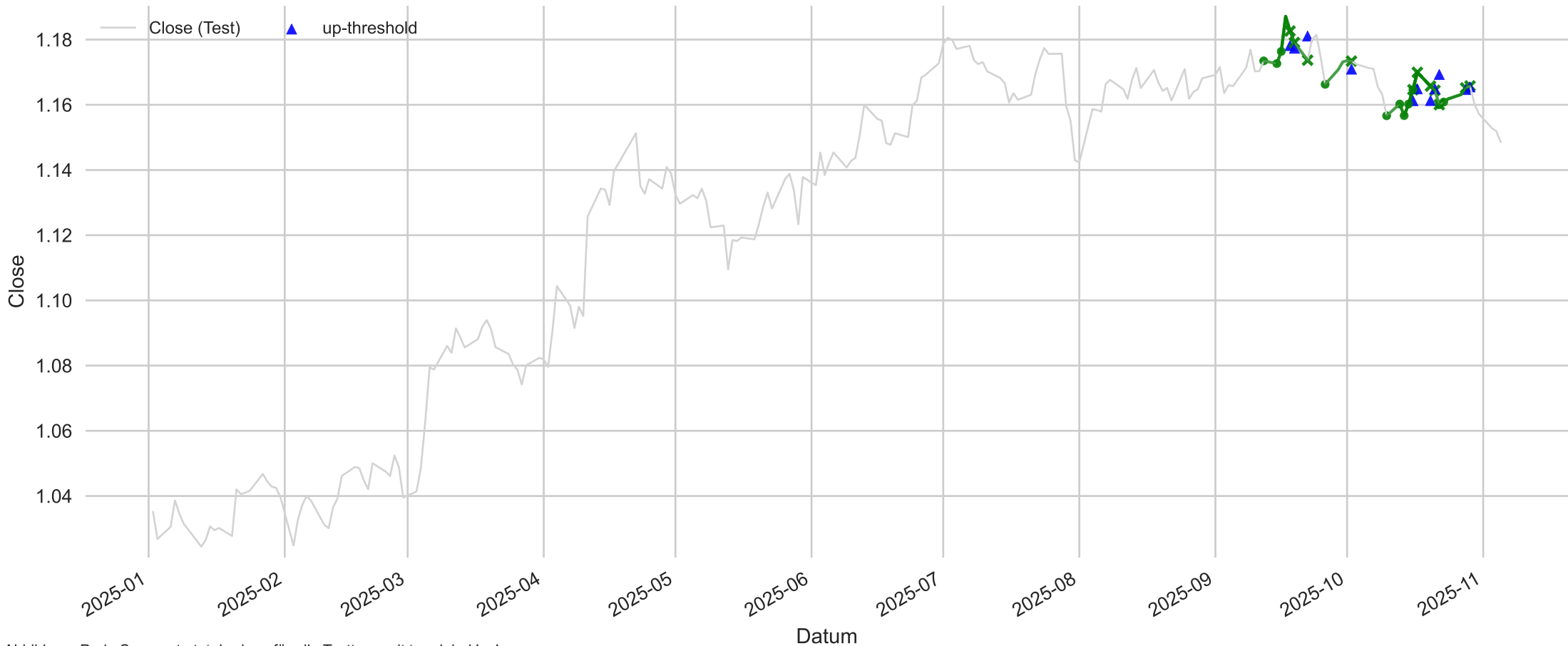


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 1

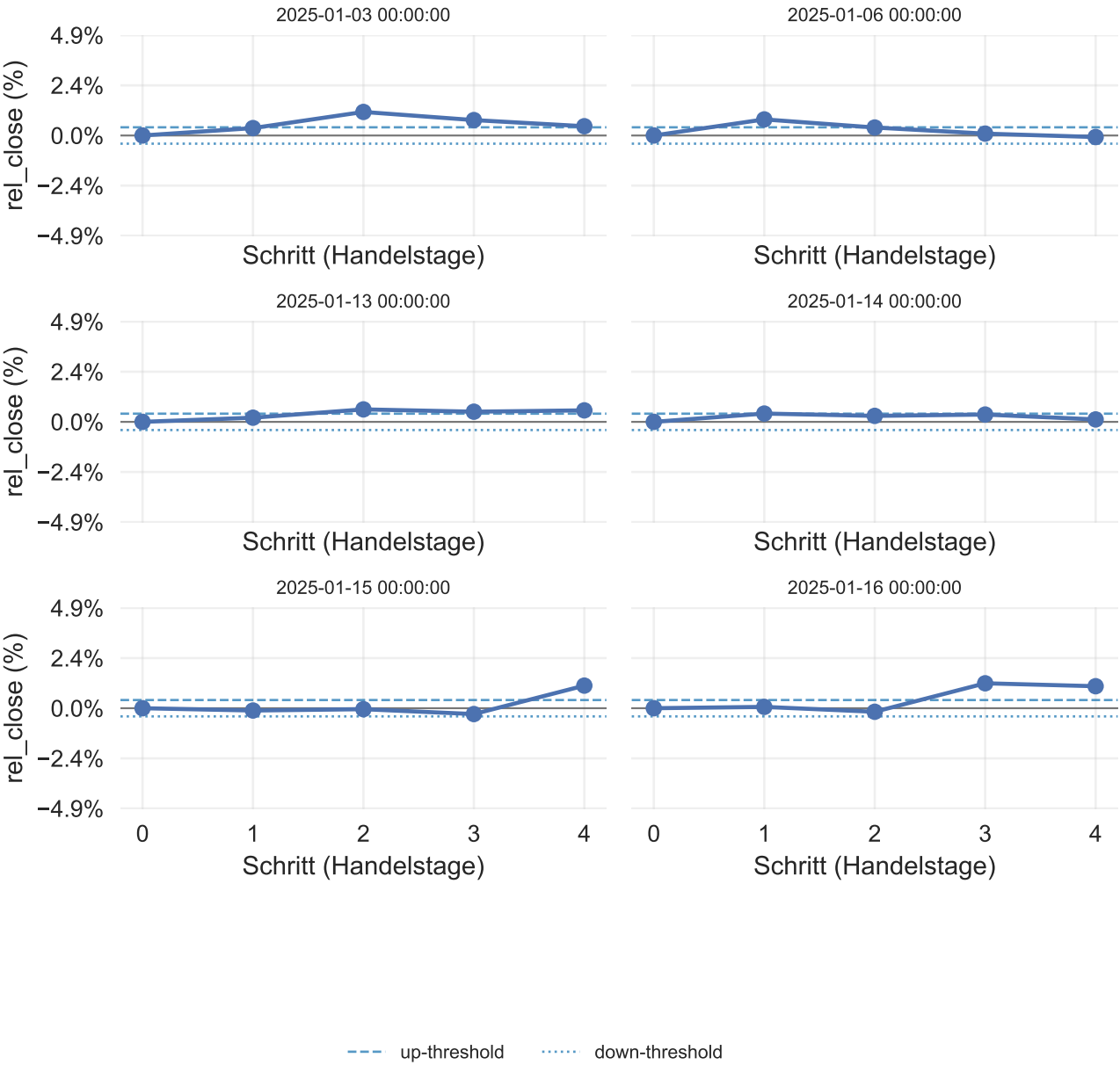


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 2

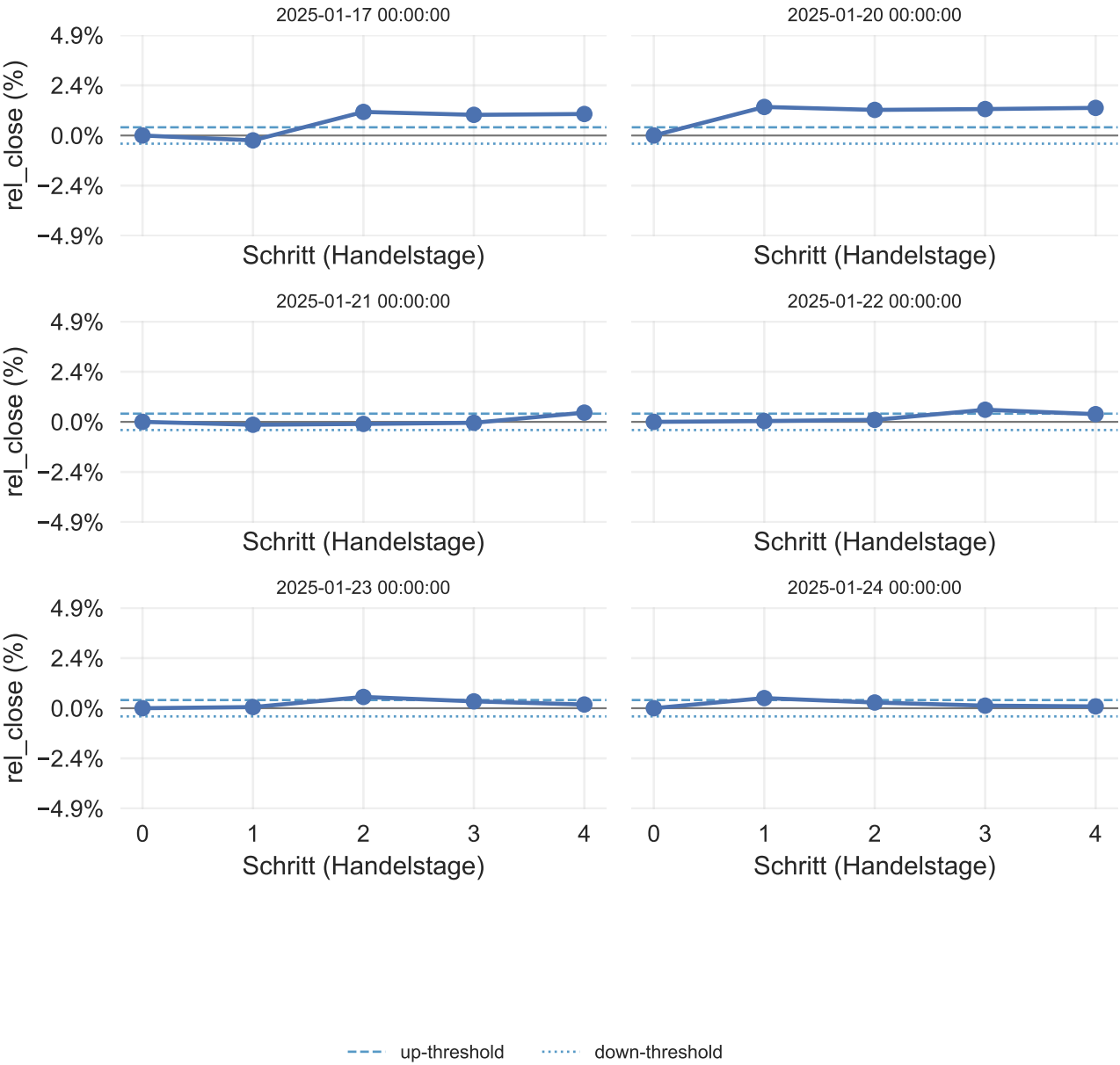


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.



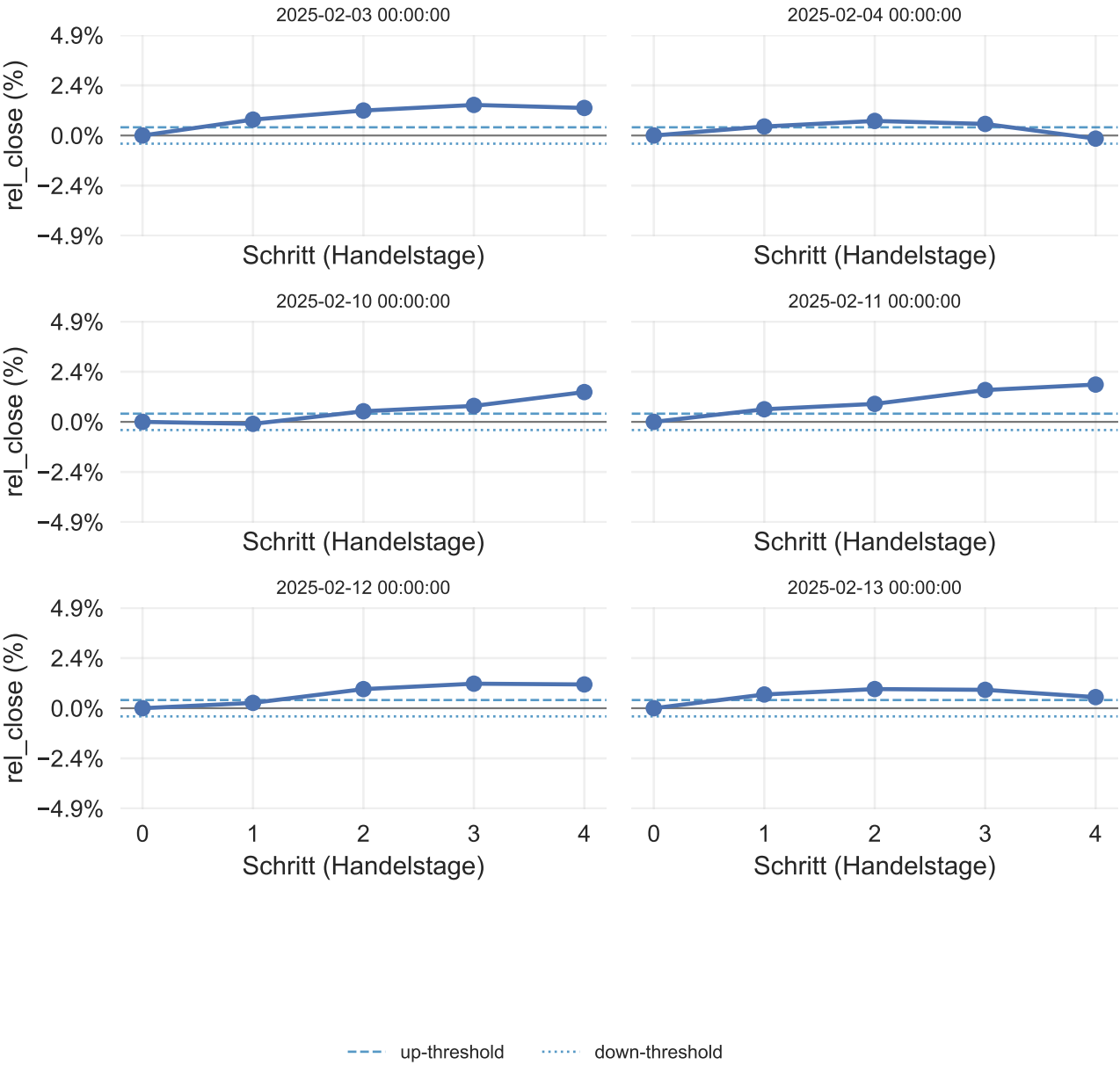


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 4

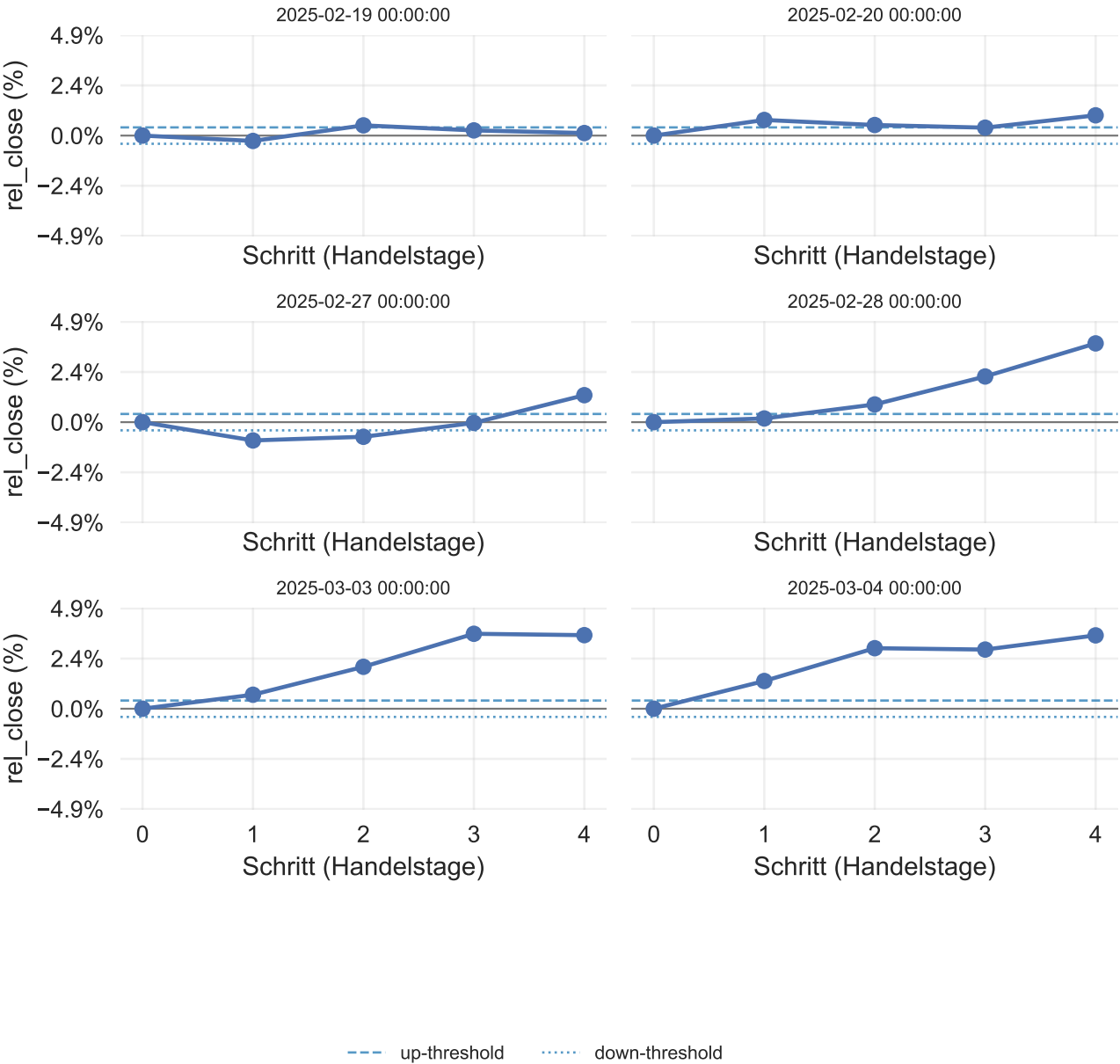


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

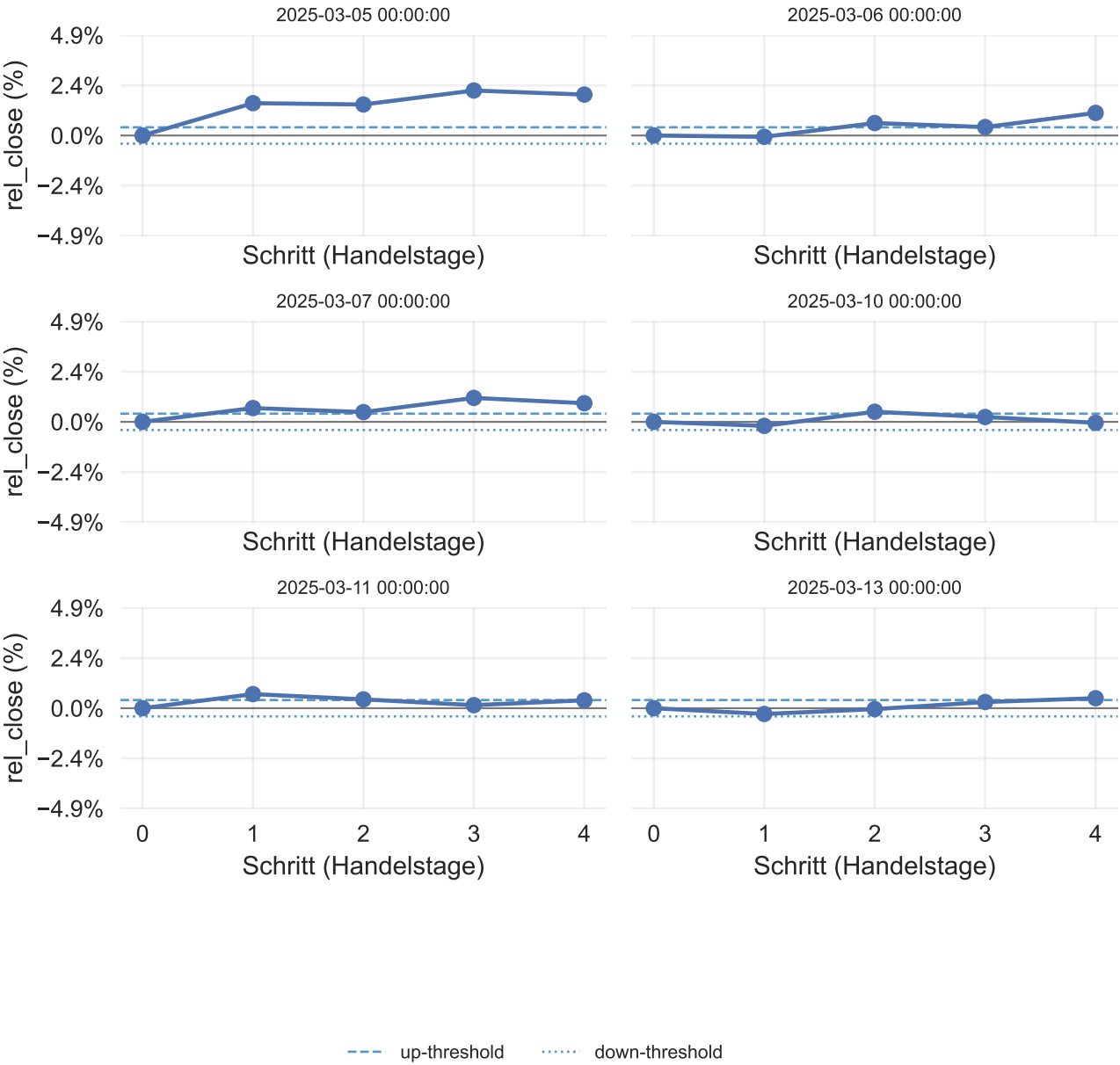


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

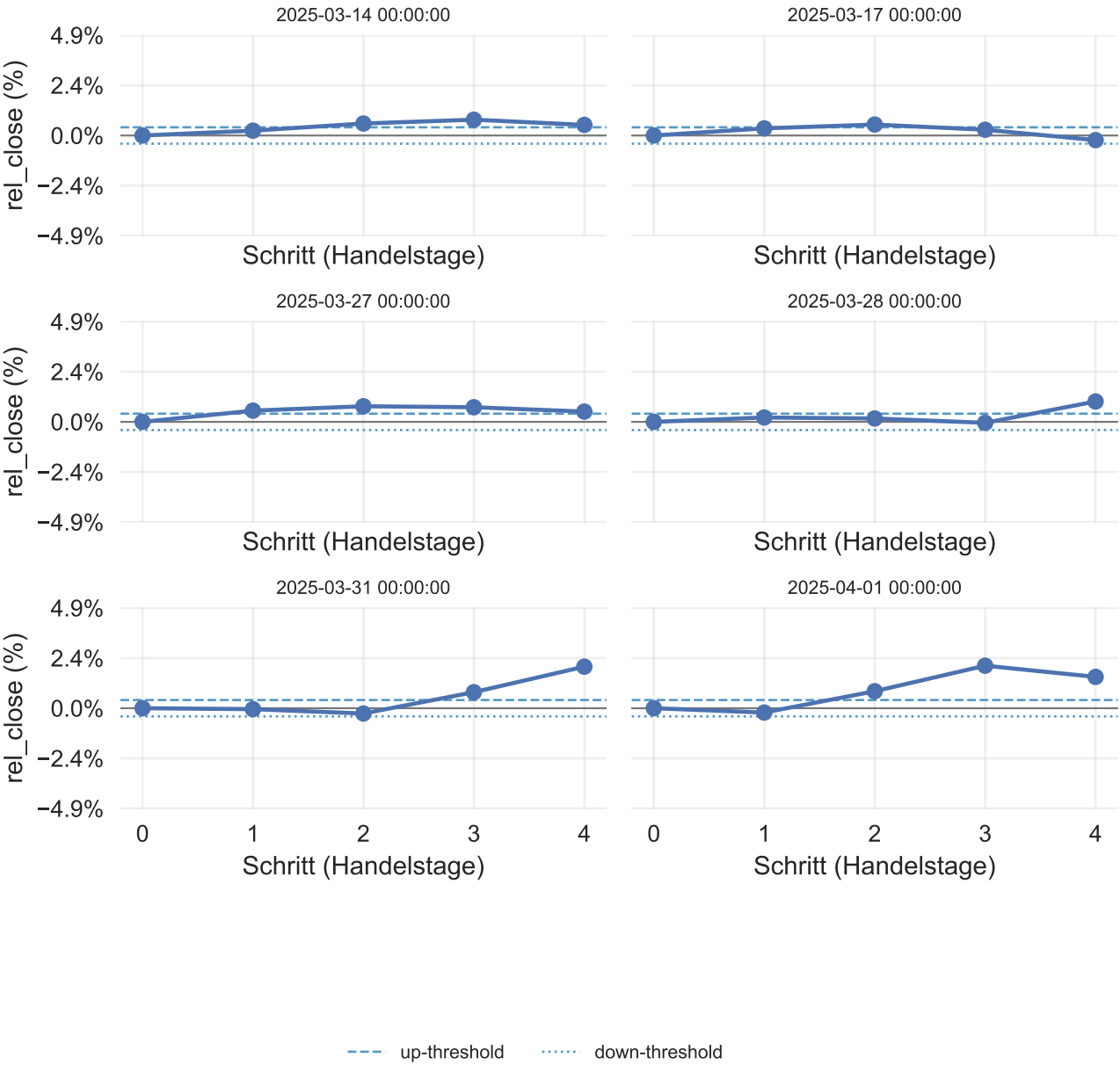


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

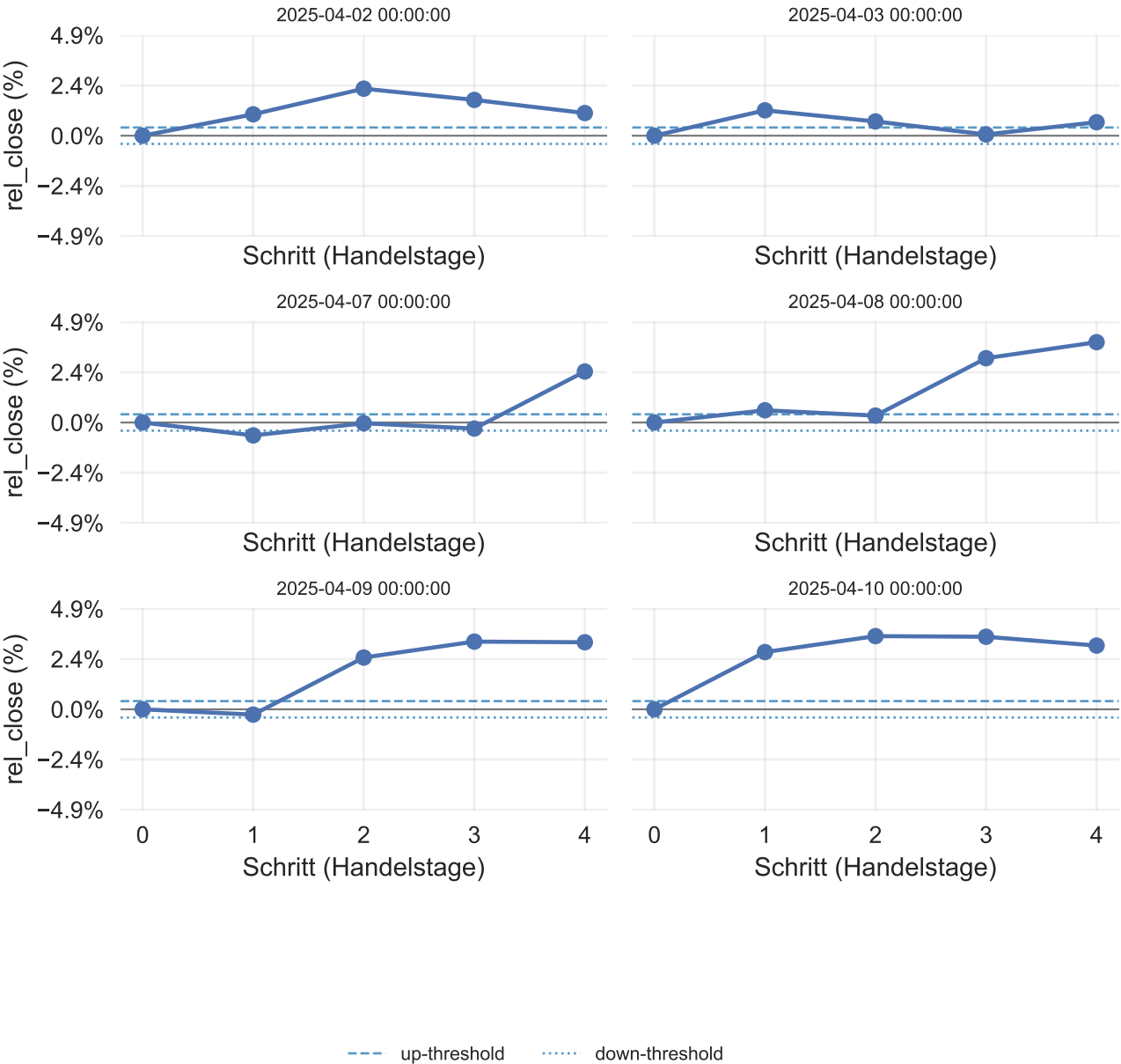


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

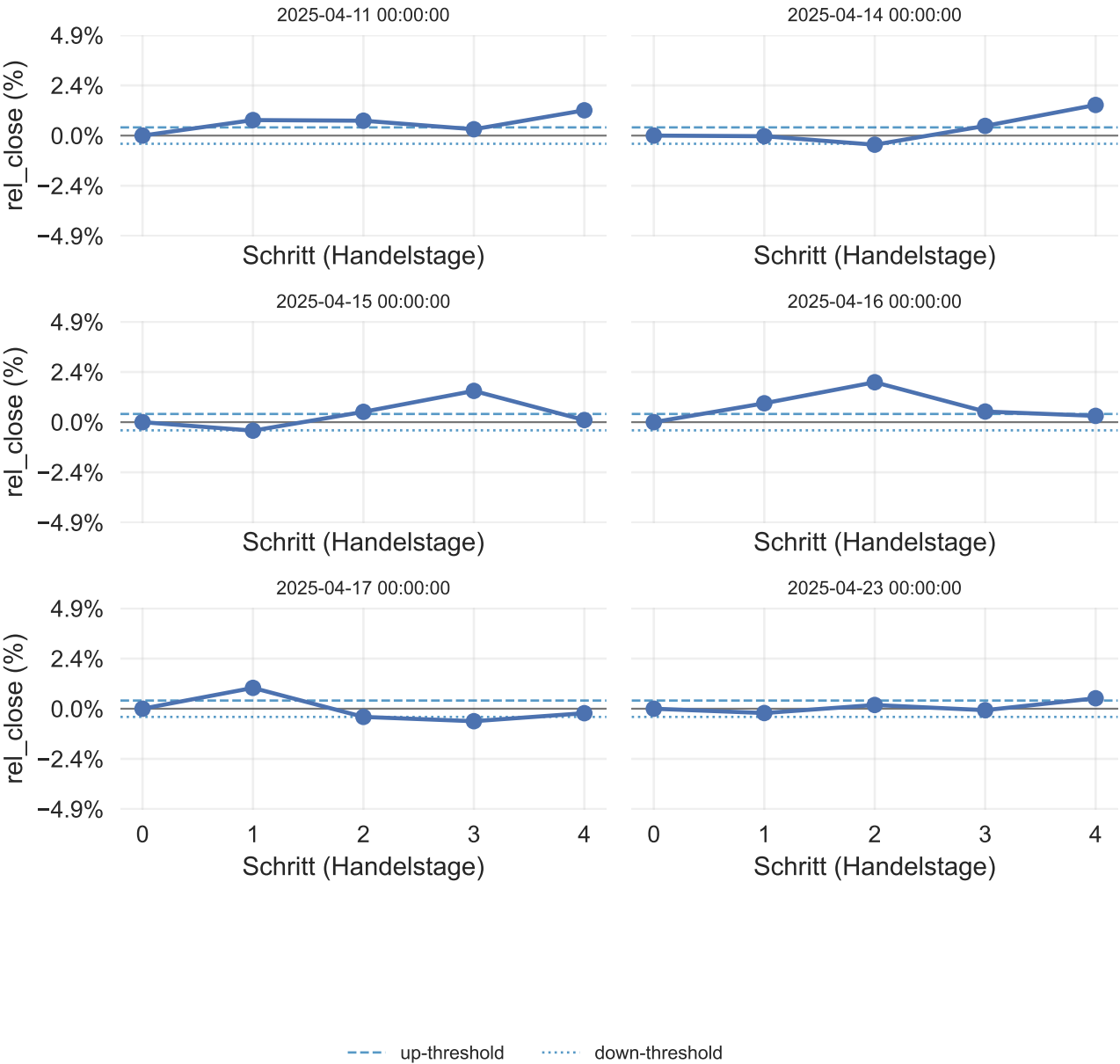


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

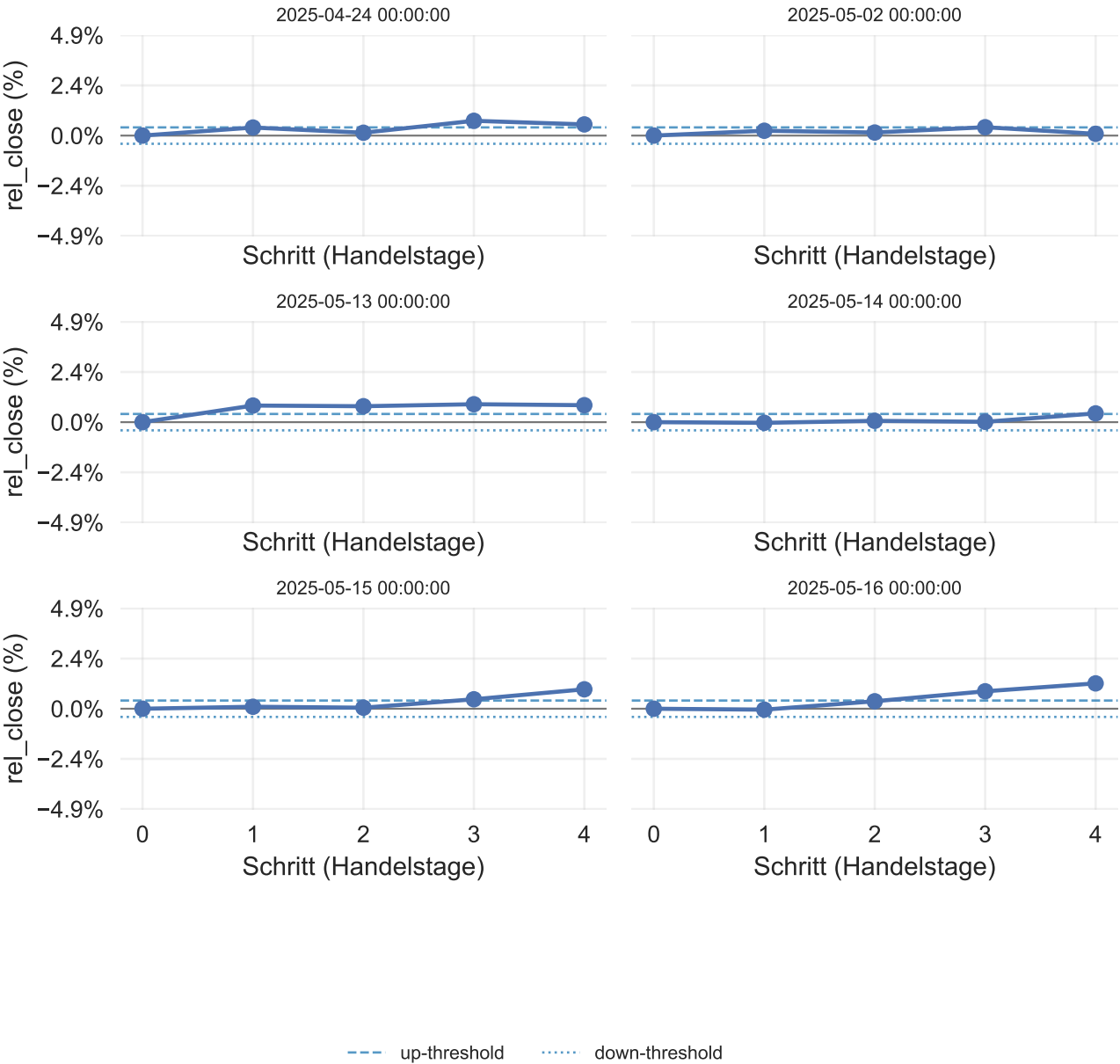


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

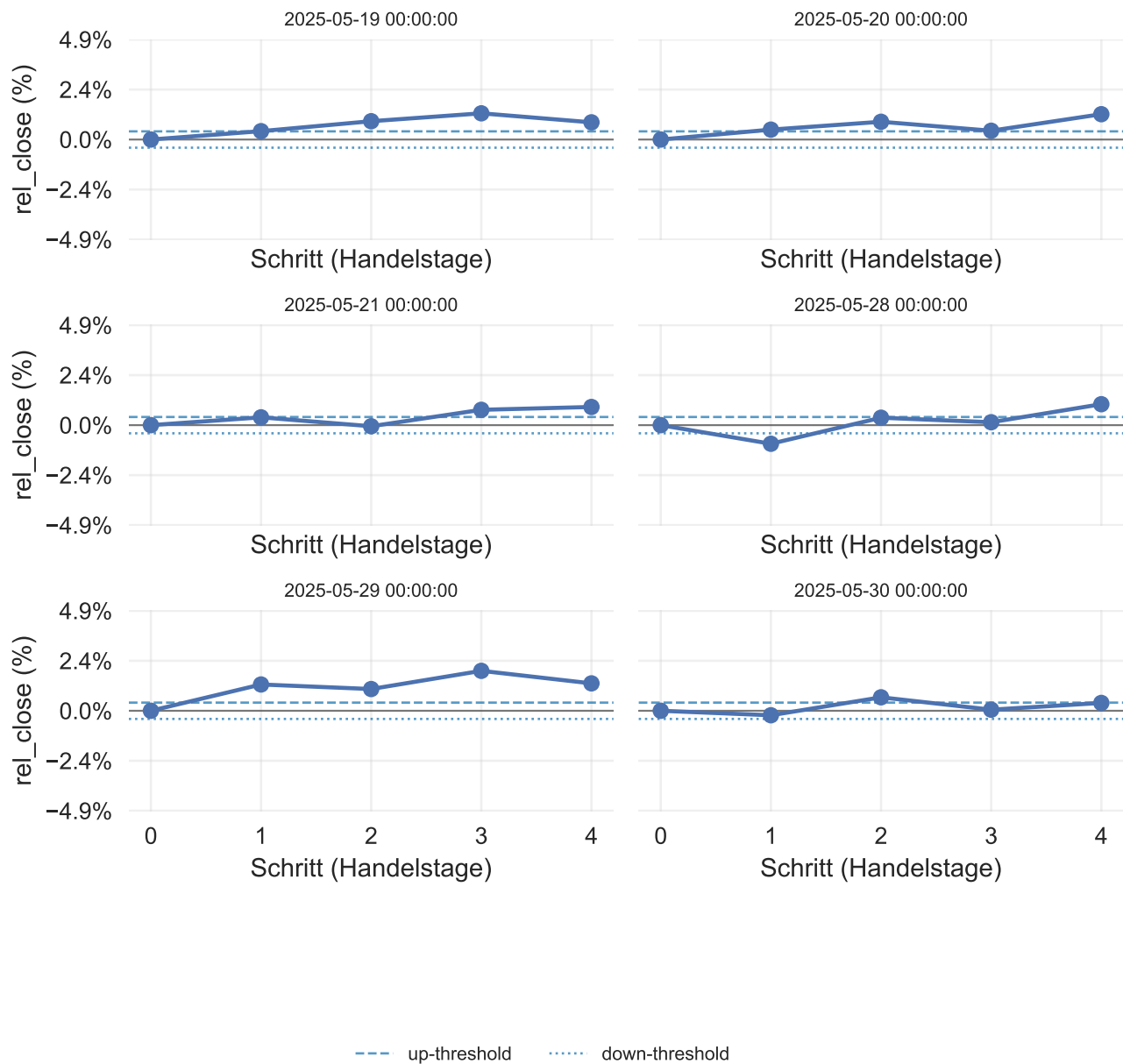


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.



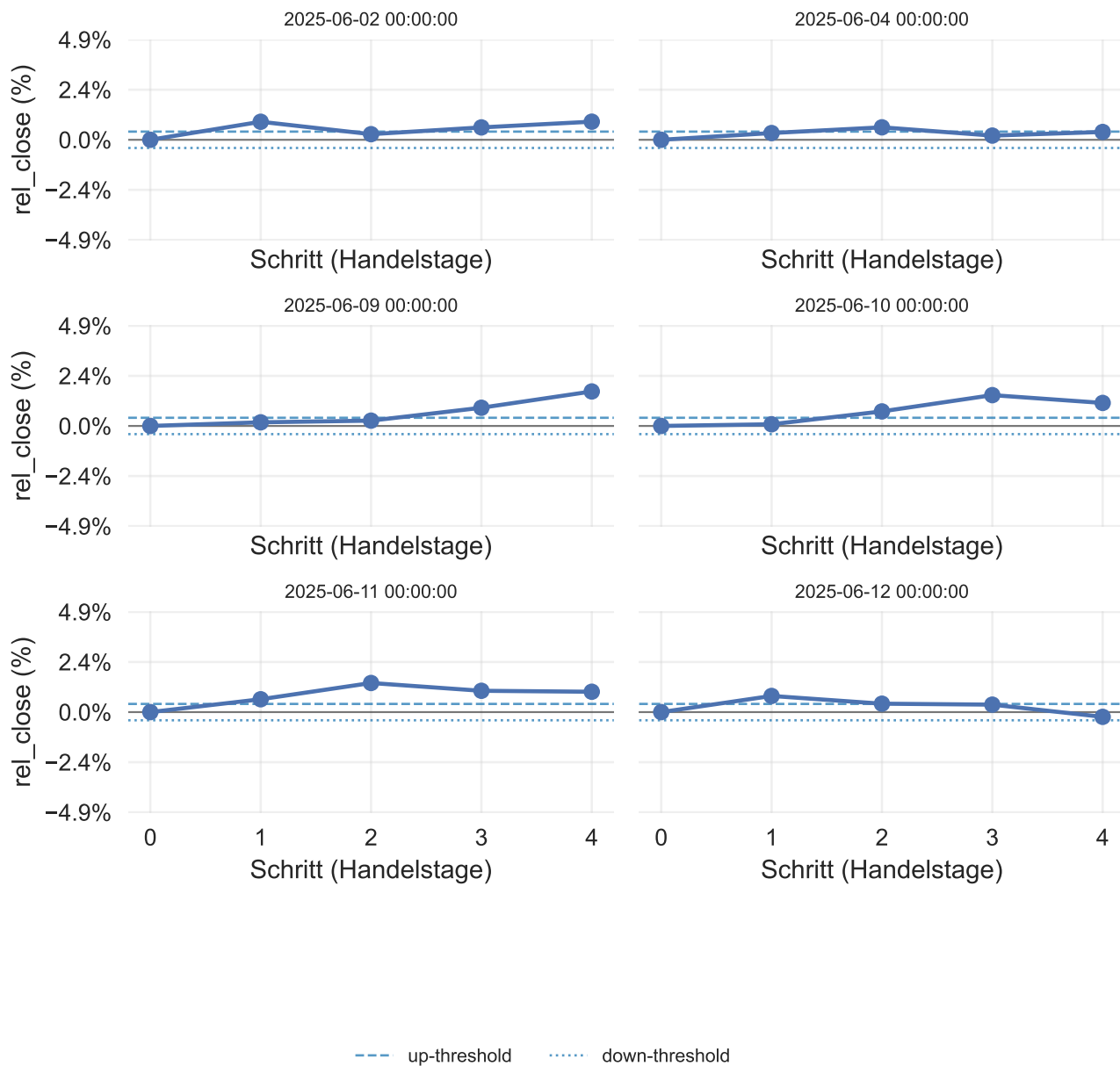


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

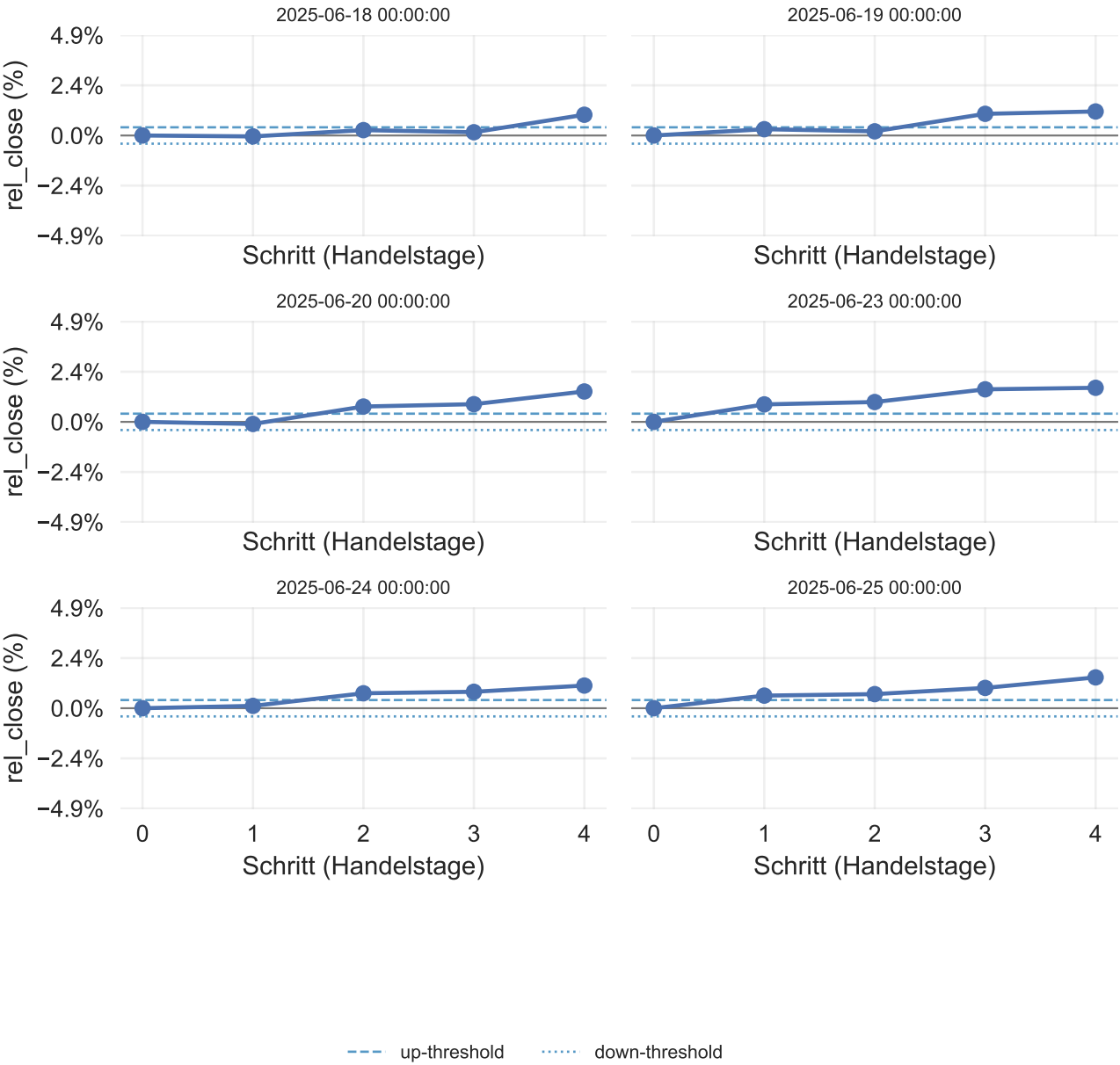


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

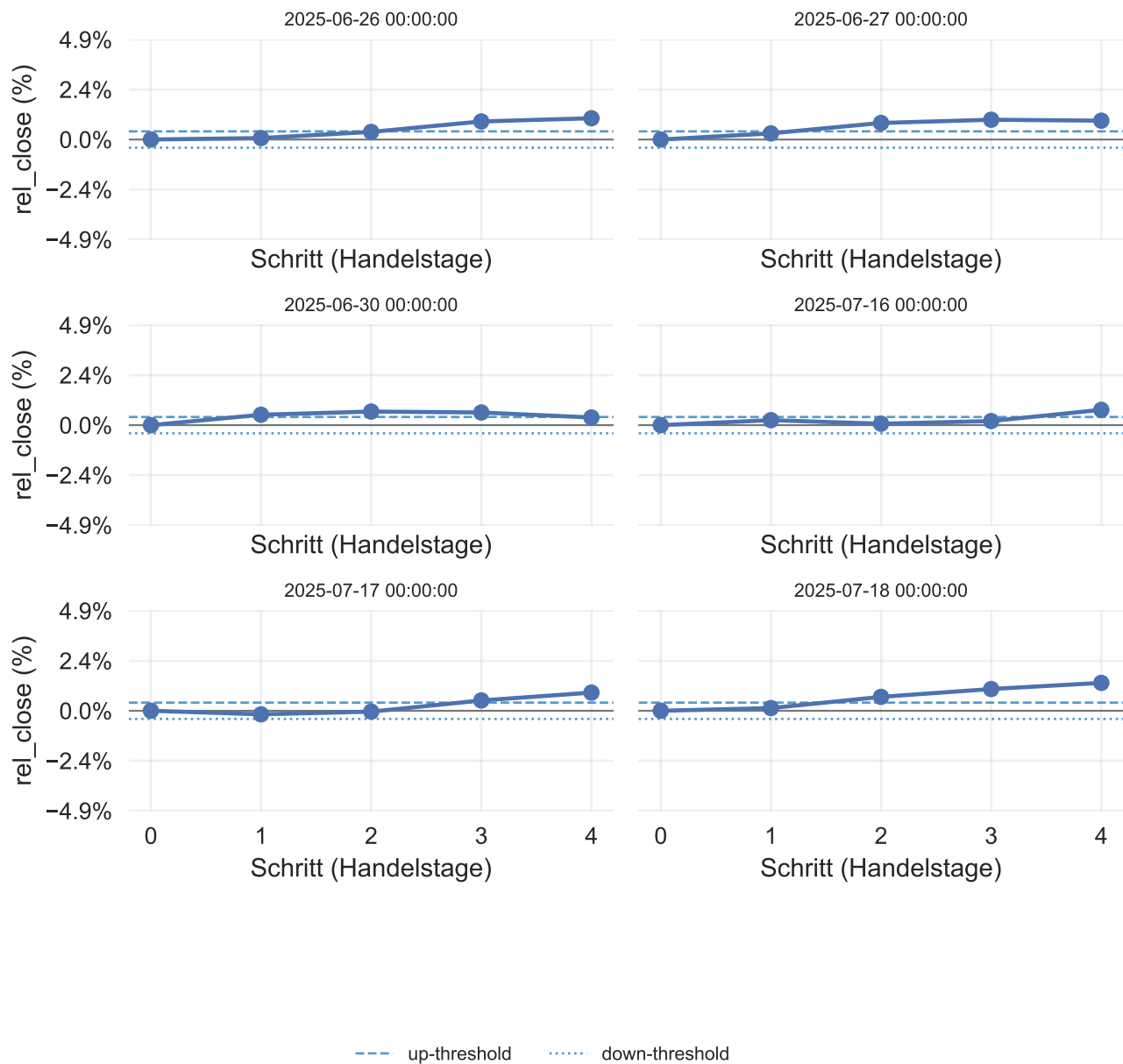


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

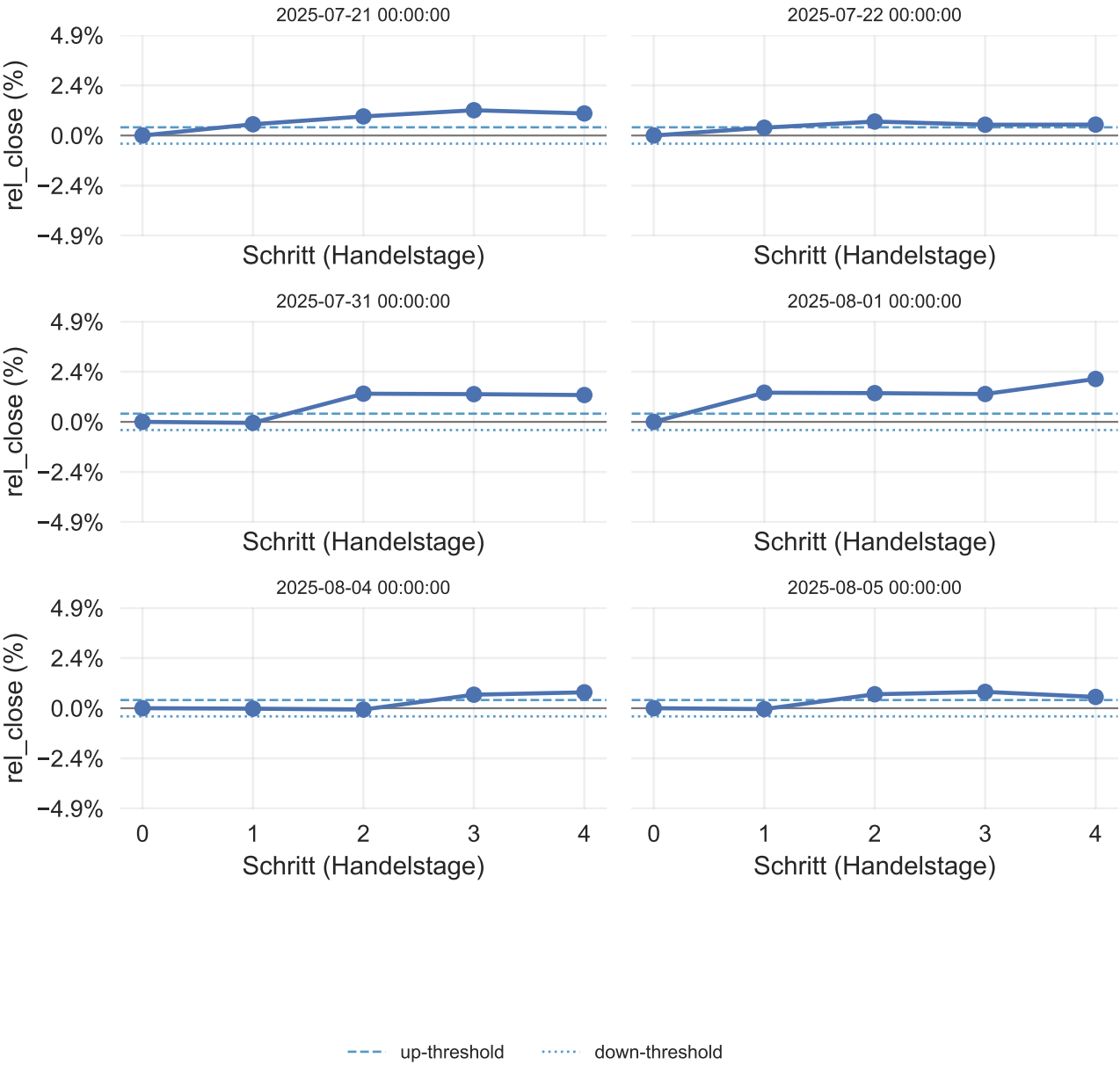


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

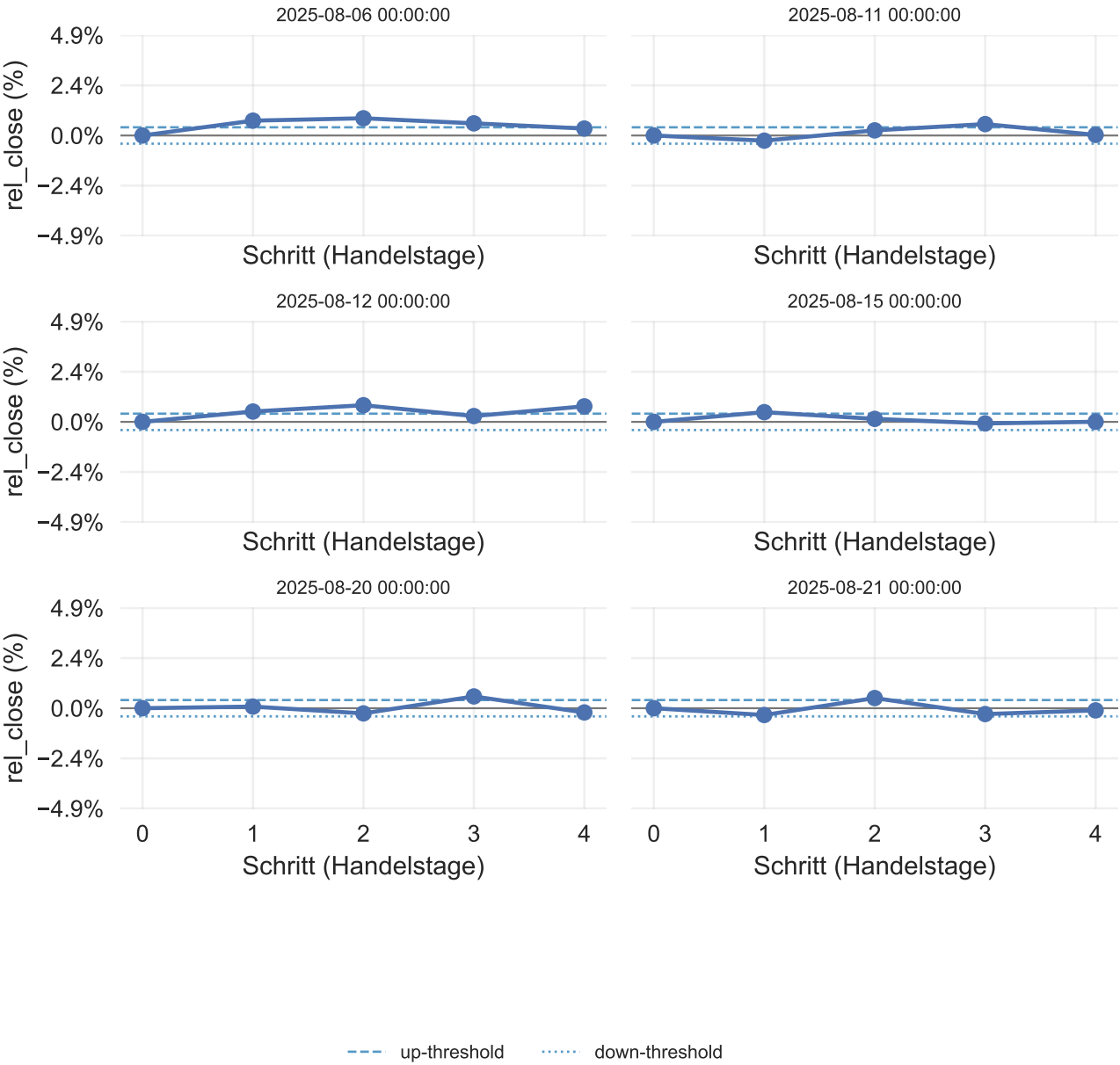


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

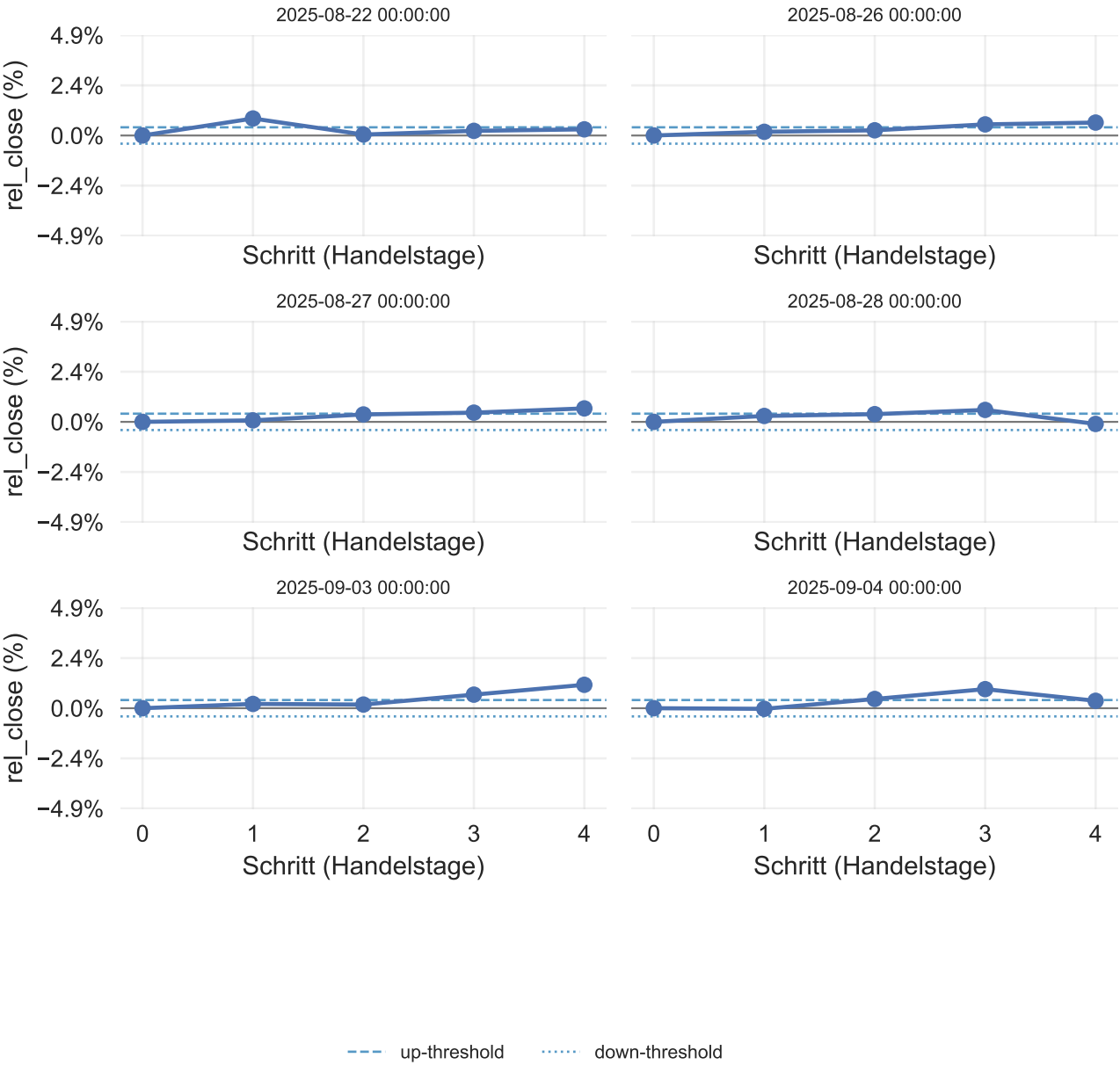


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

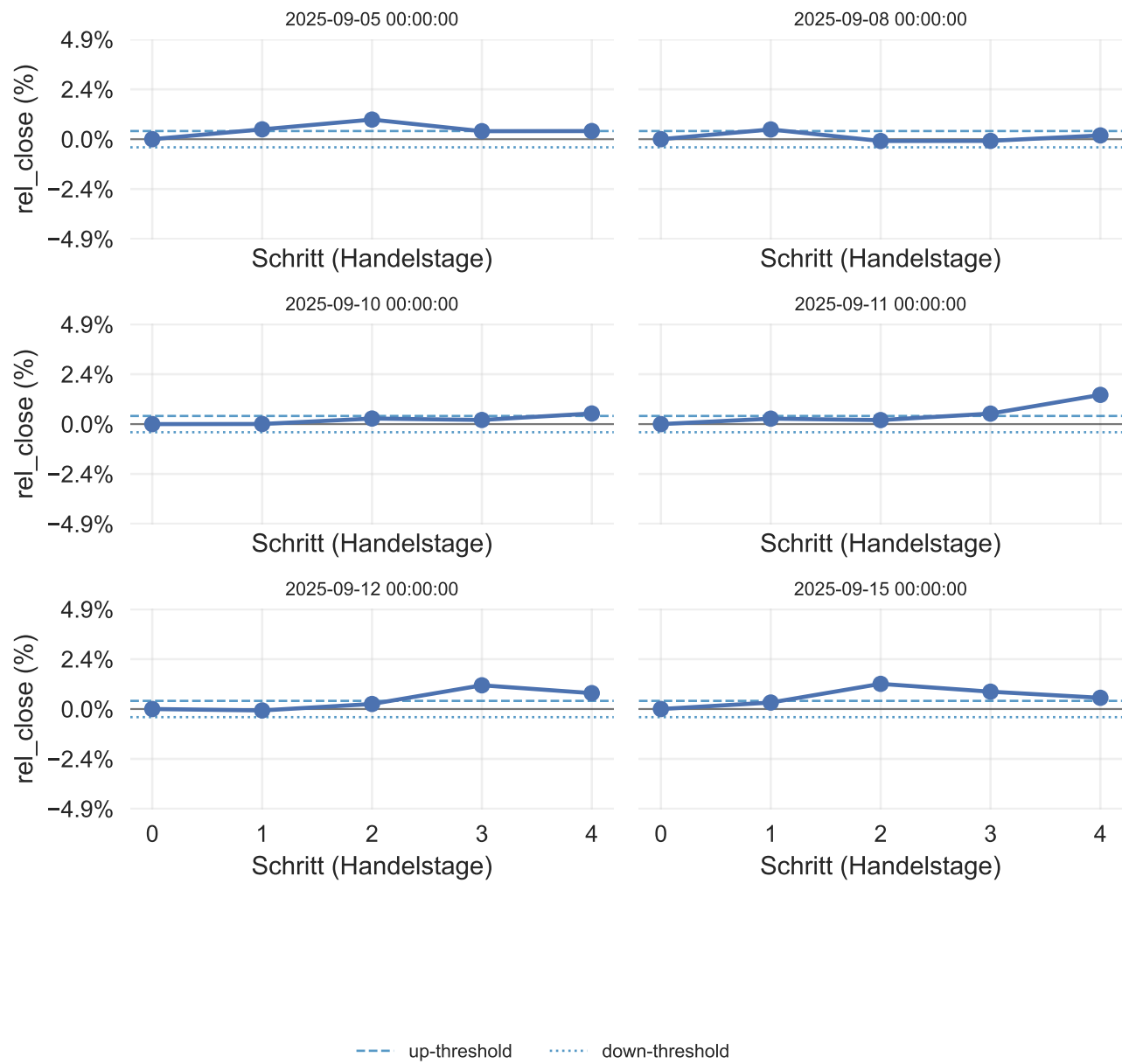


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

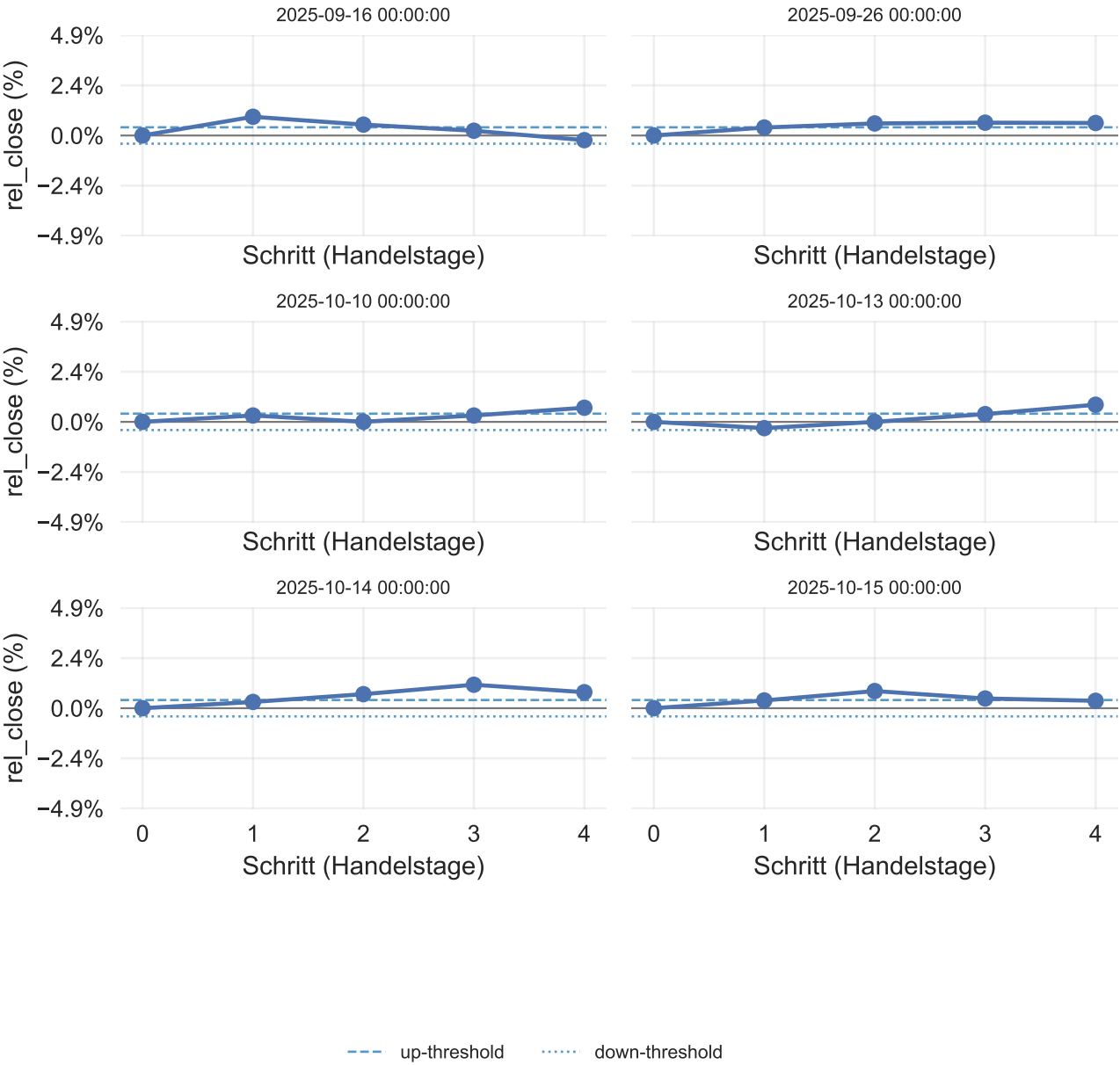


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.



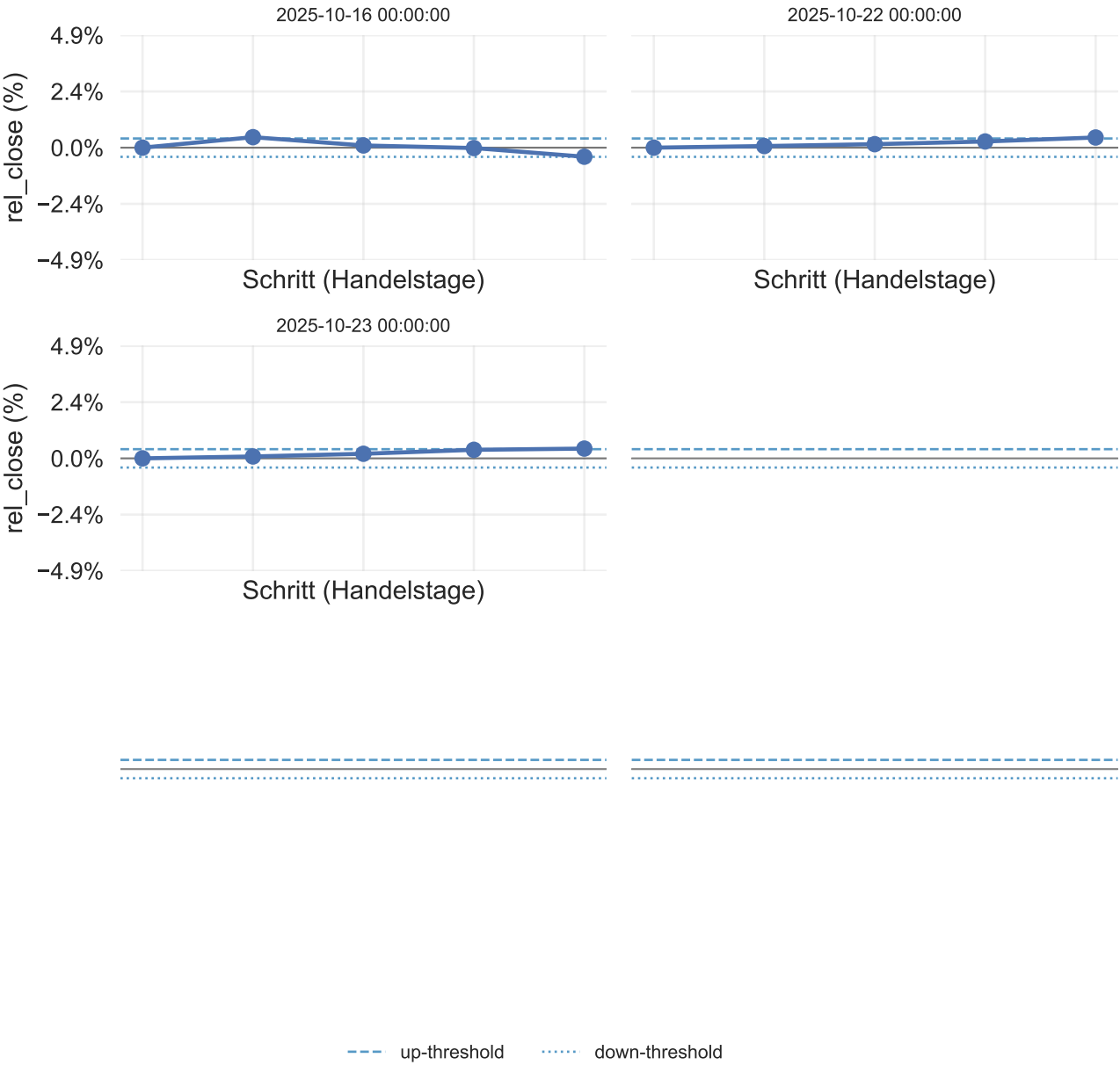


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

# EURUSD-Segmente mit label='down' (Test-Split) – Seite 1/4

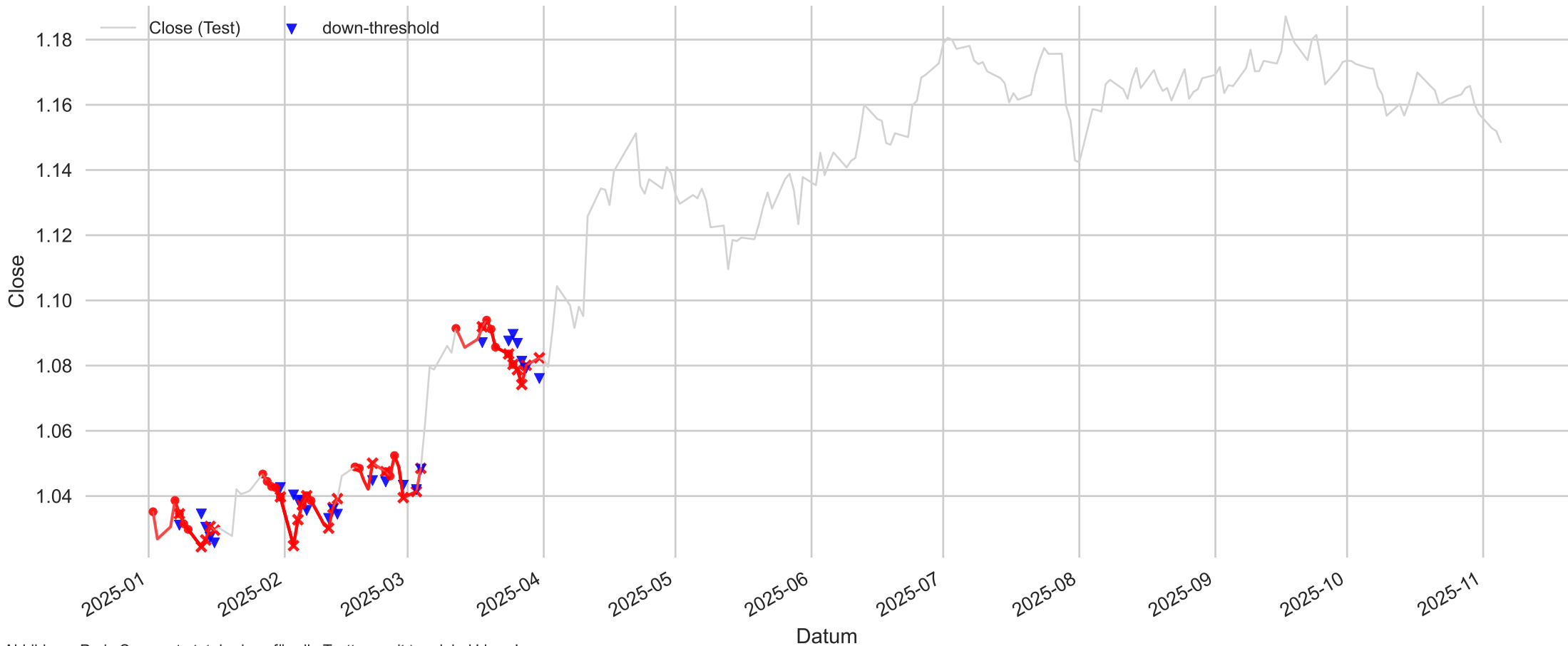


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'down'.

# EURUSD-Segmente mit label='down' (Test-Split) – Seite 2/4

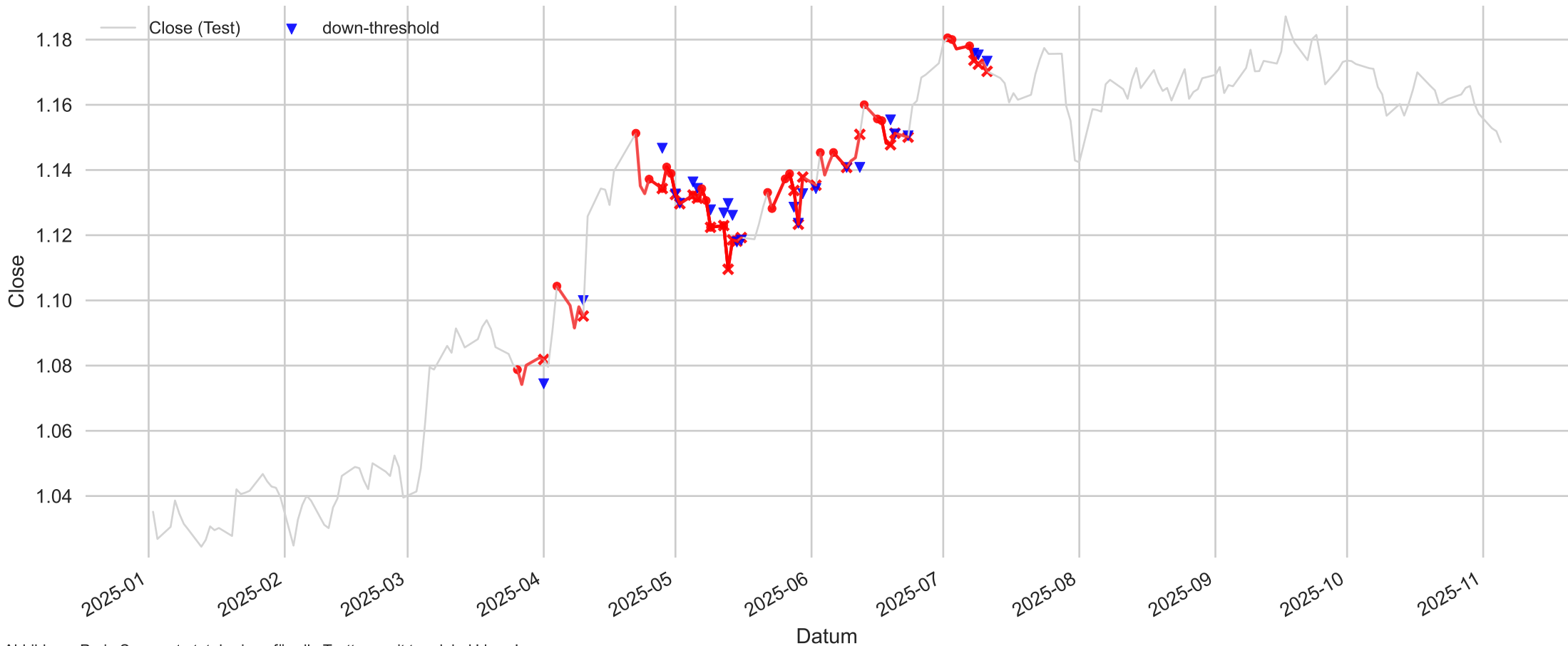


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'down'.

# EURUSD-Segmente mit label='down' (Test-Split) – Seite 3/4

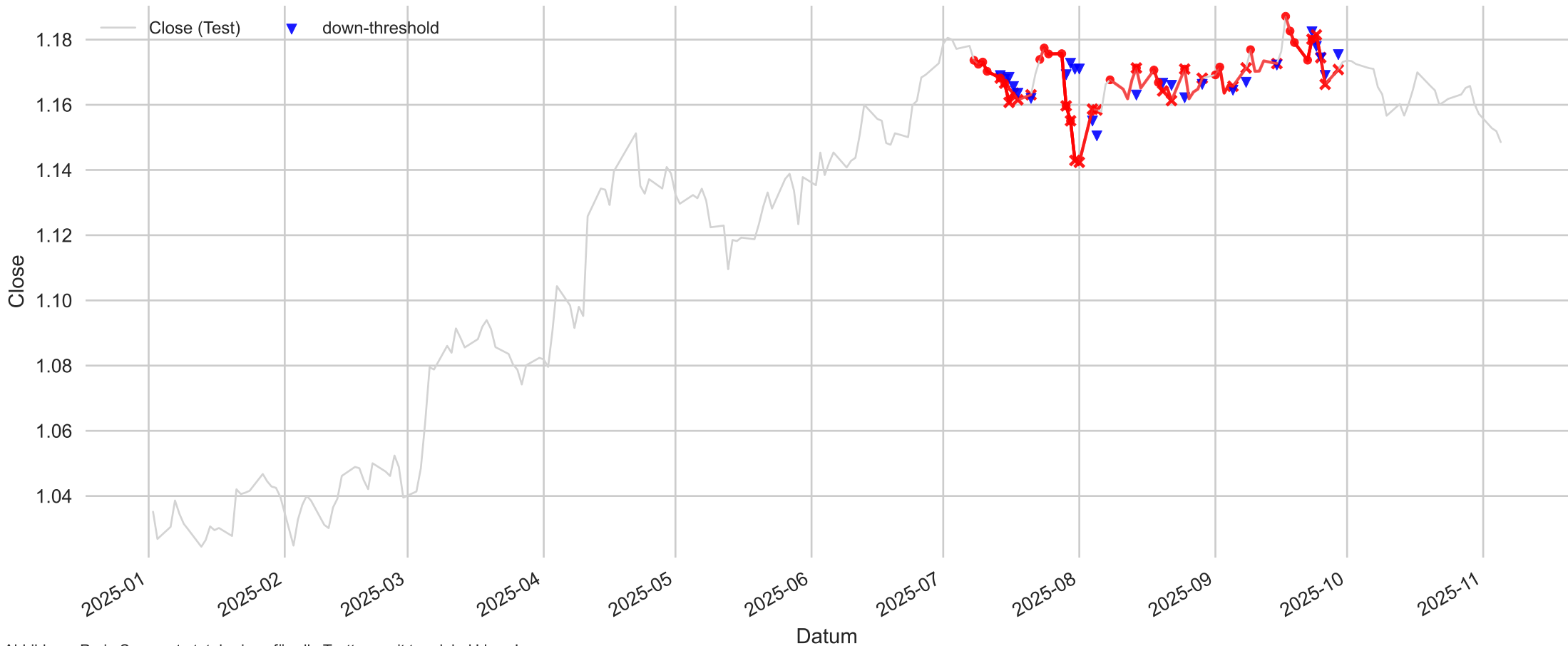


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'down'.

# EURUSD-Segmente mit label='down' (Test-Split) – Seite 4/4

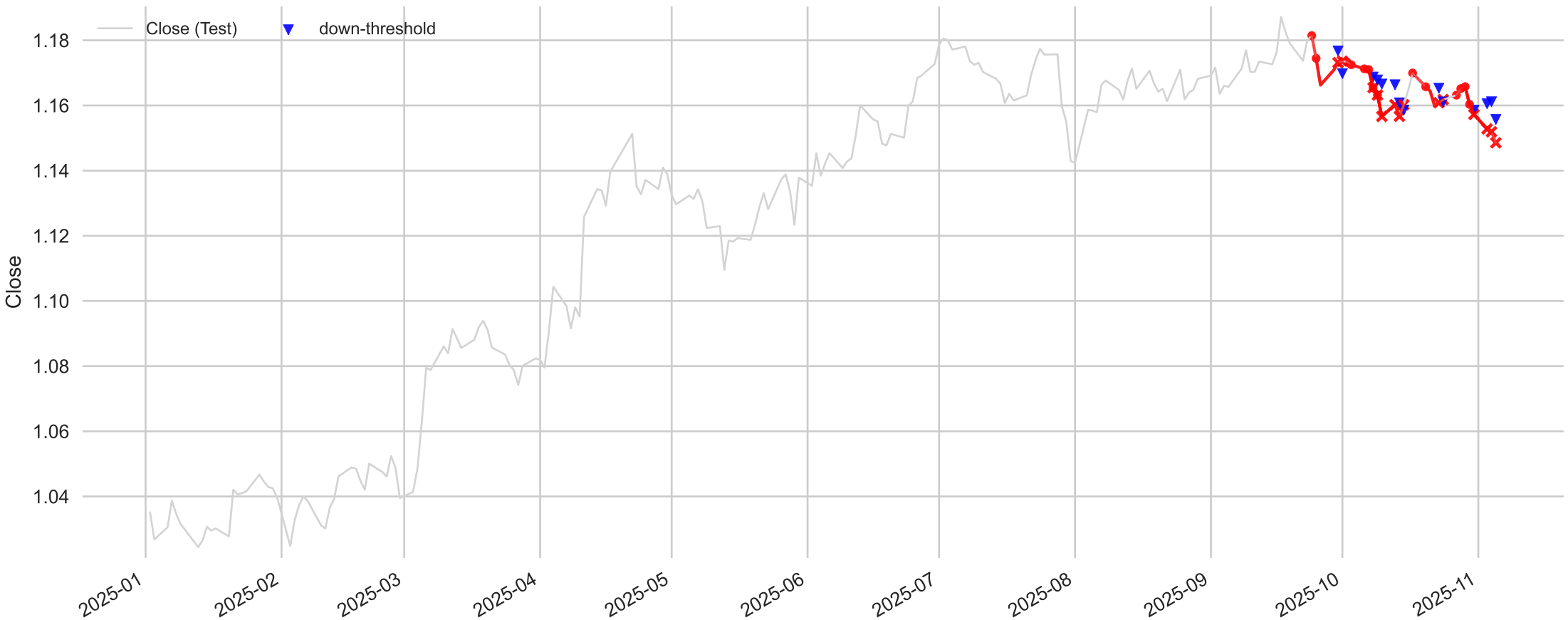


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'down'.

Relativer Verlauf der Segmente (label='down') – Seite 1

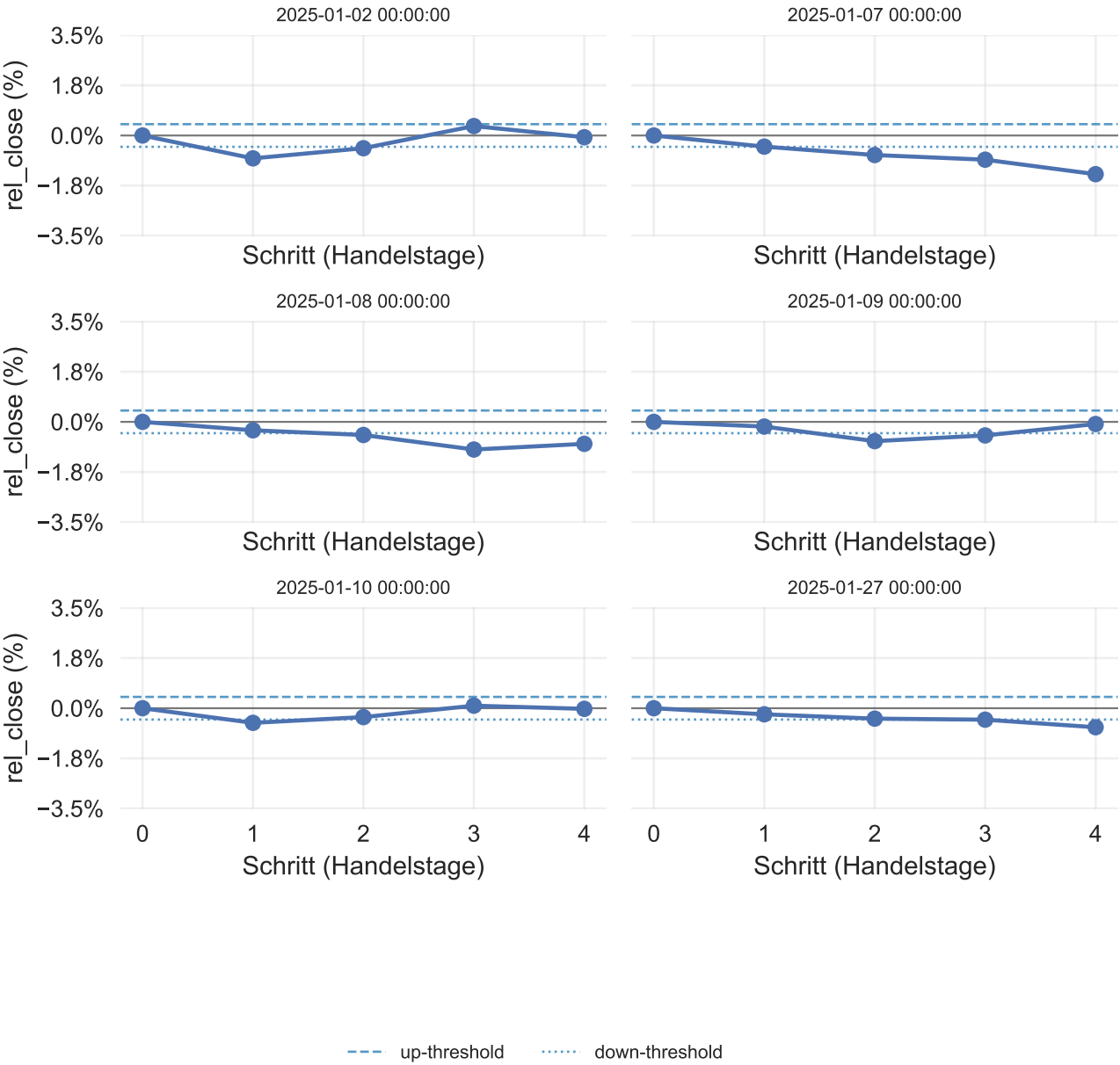


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

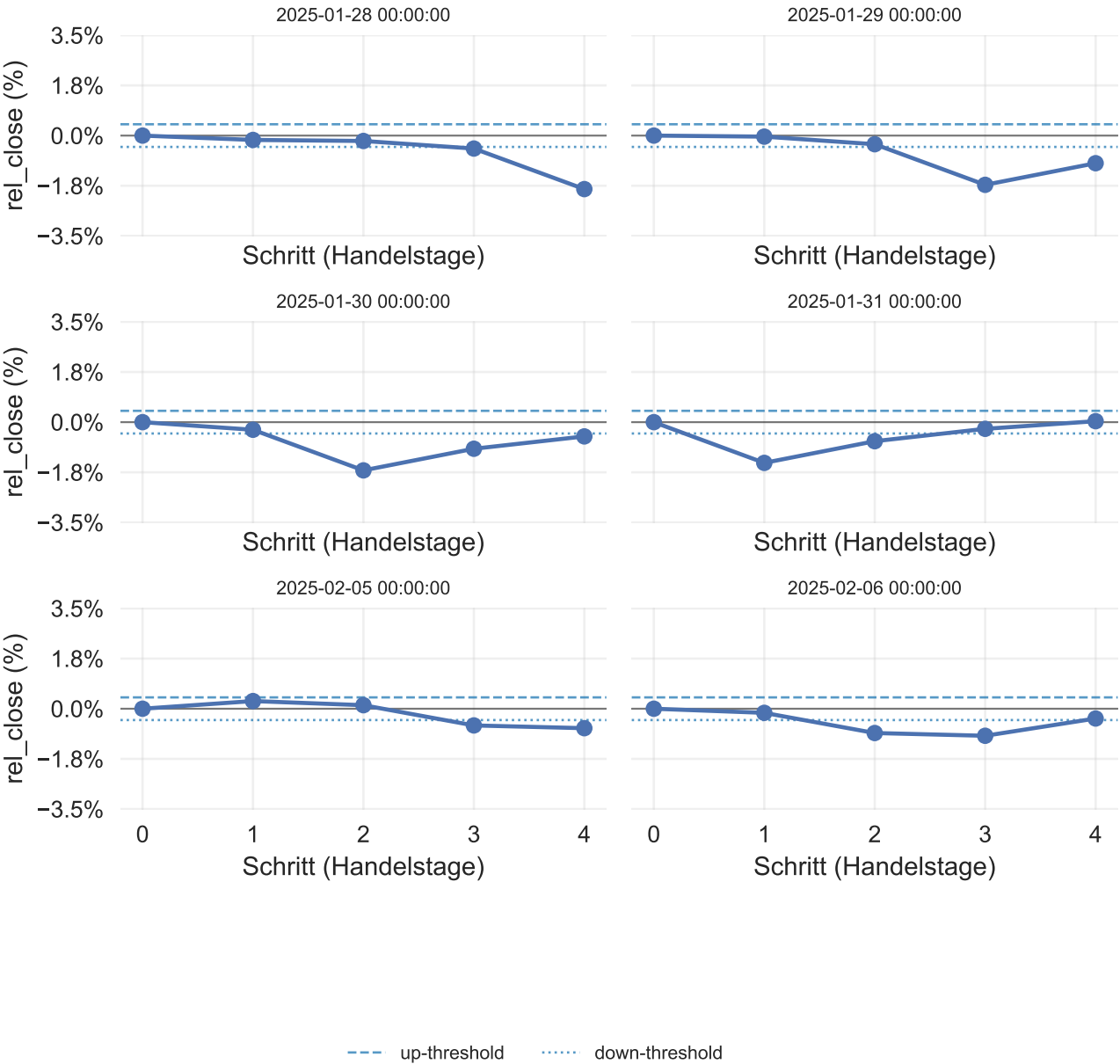


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

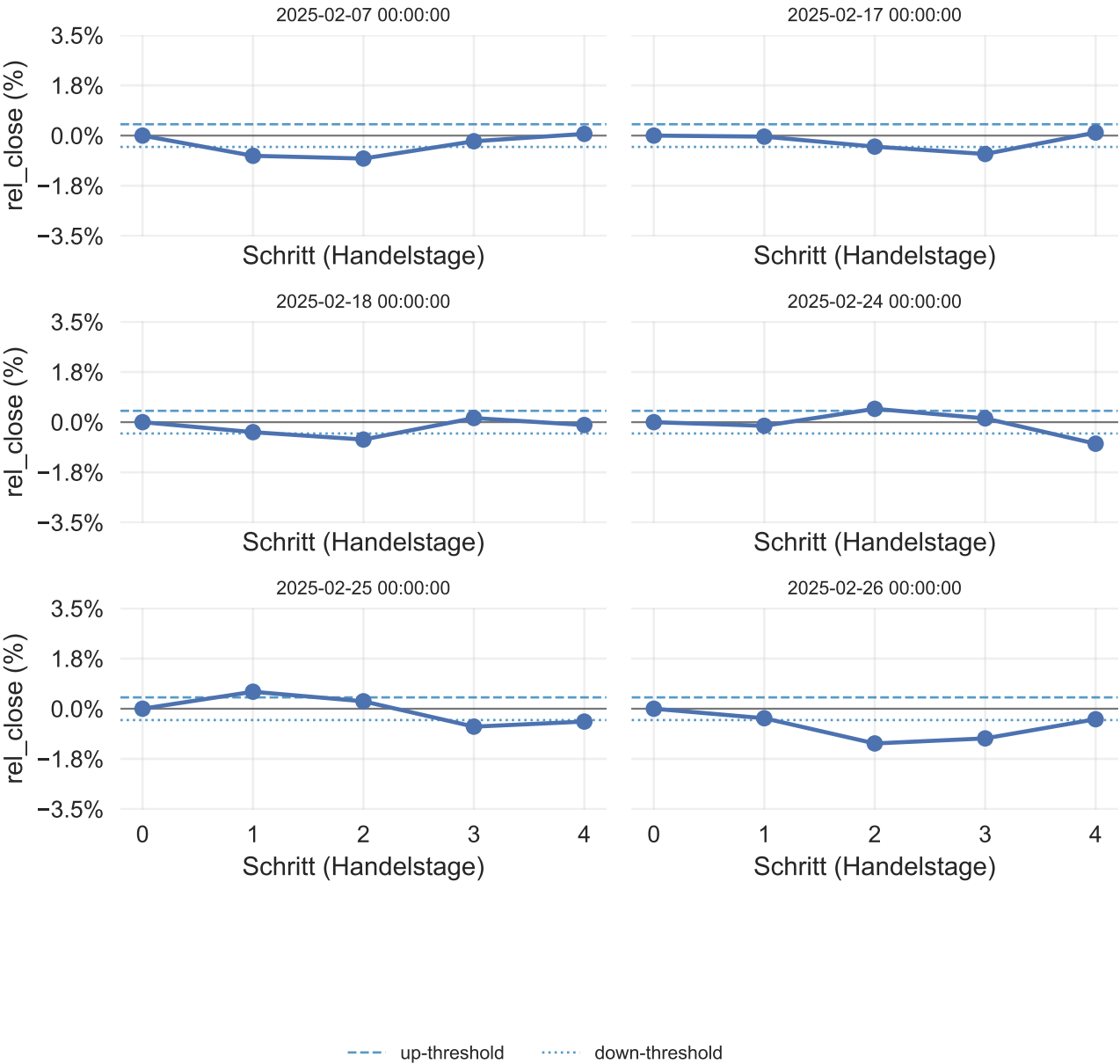


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.



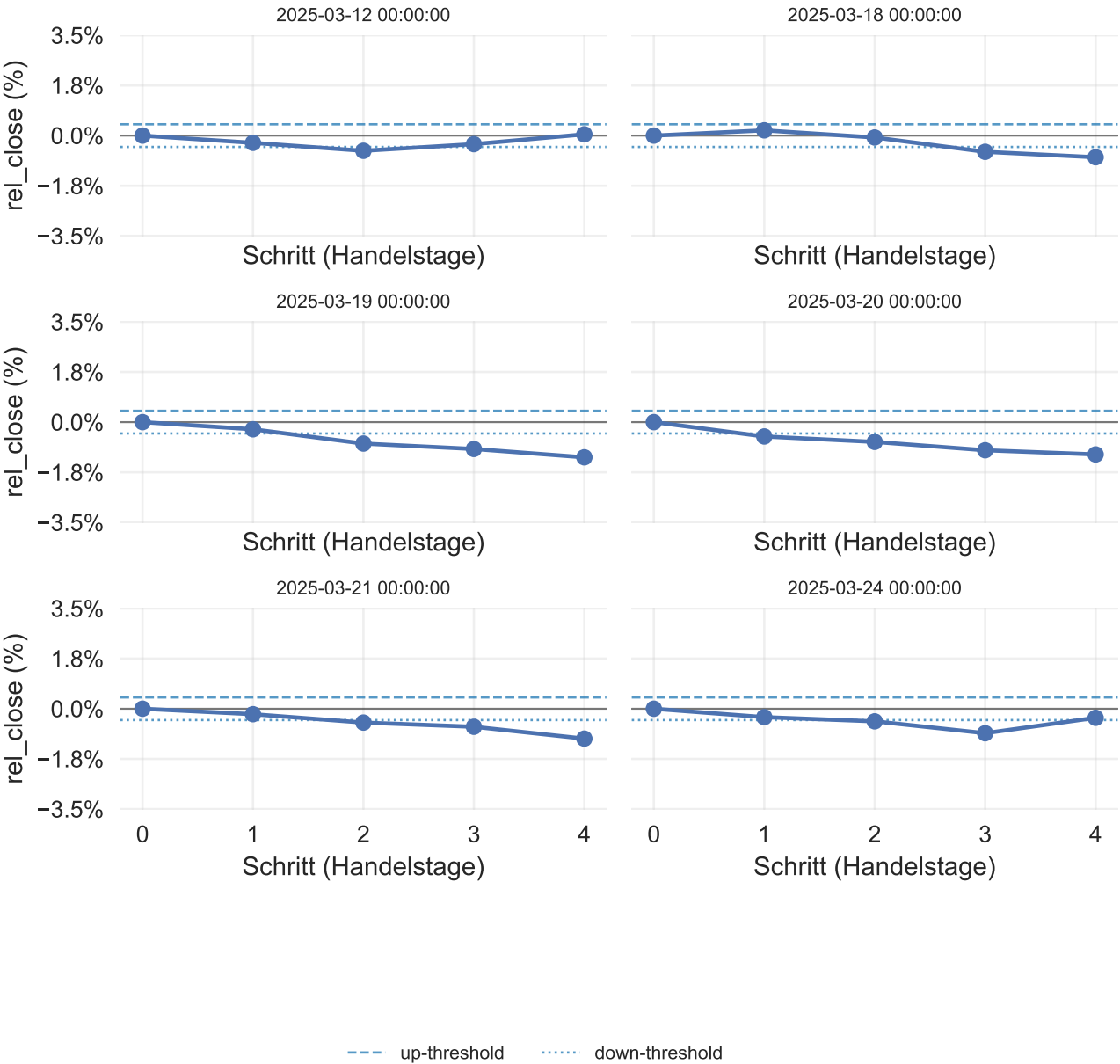


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

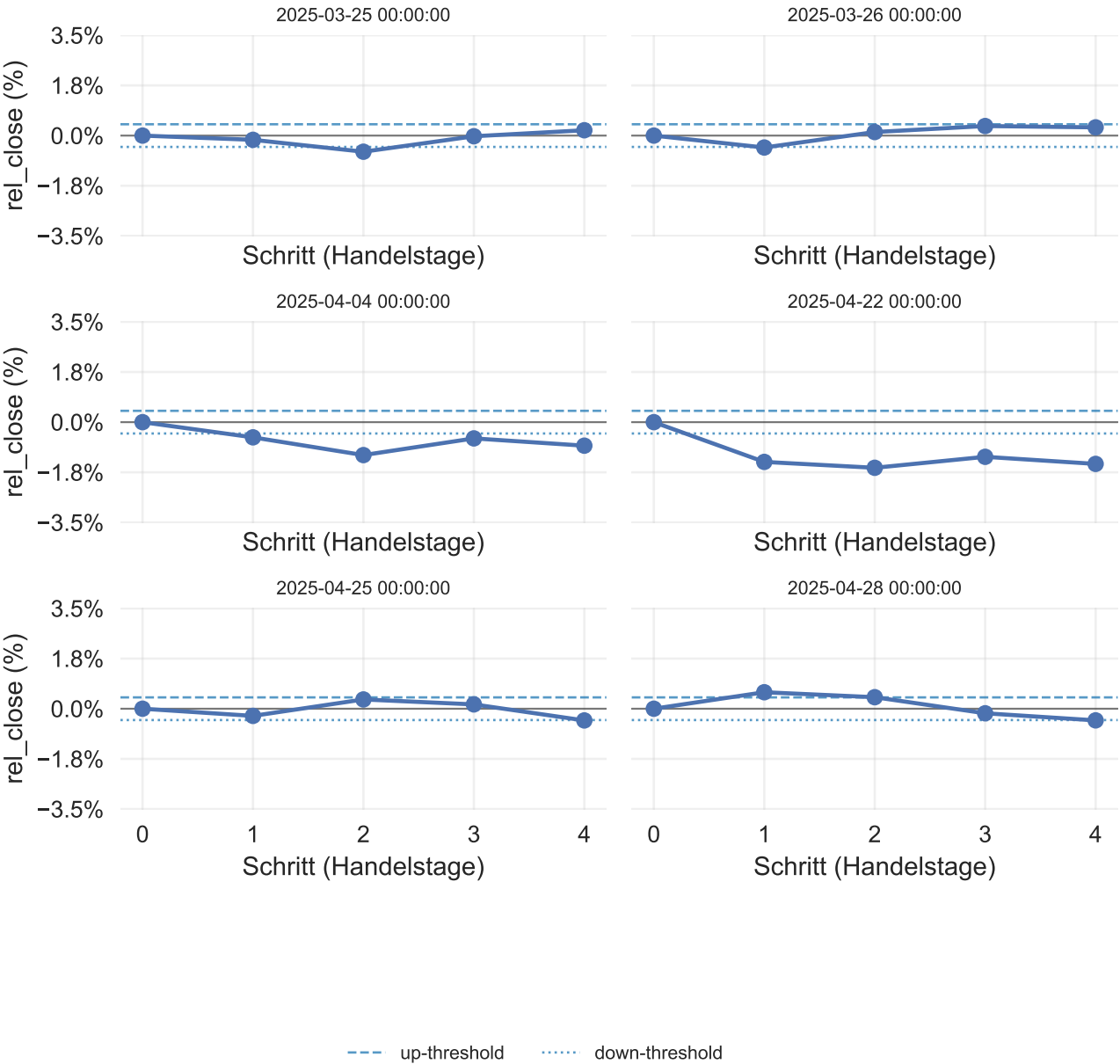


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

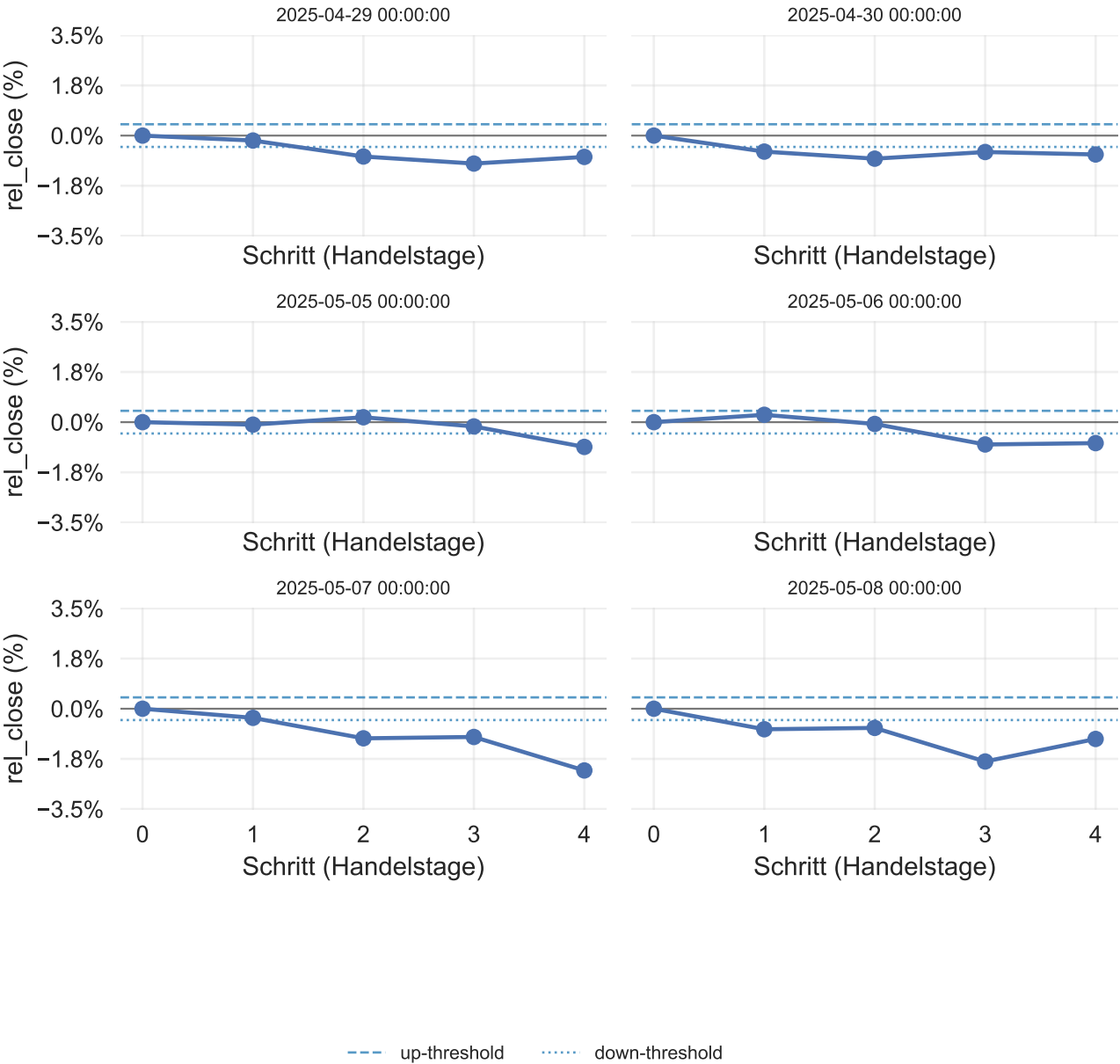


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

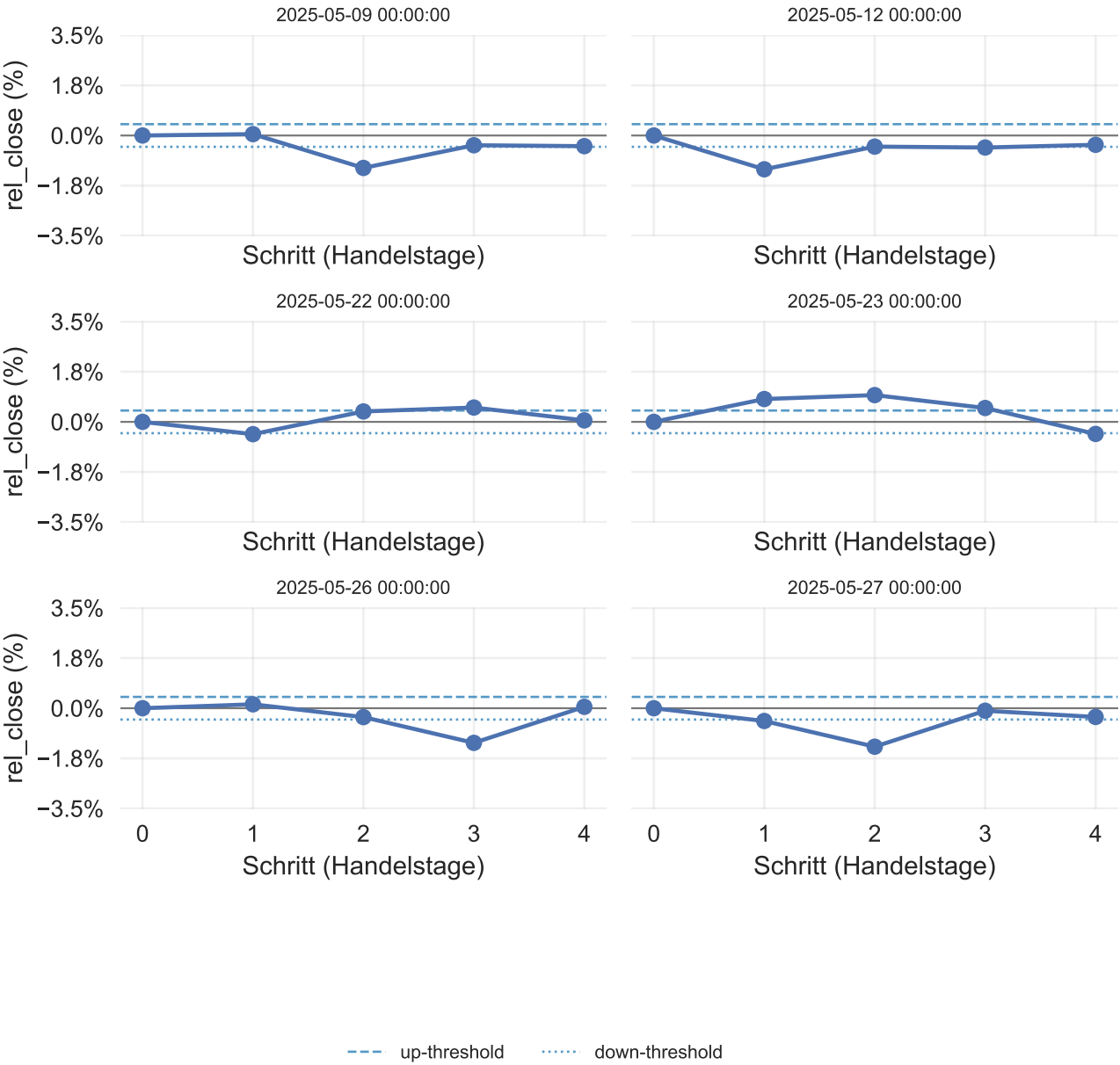


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

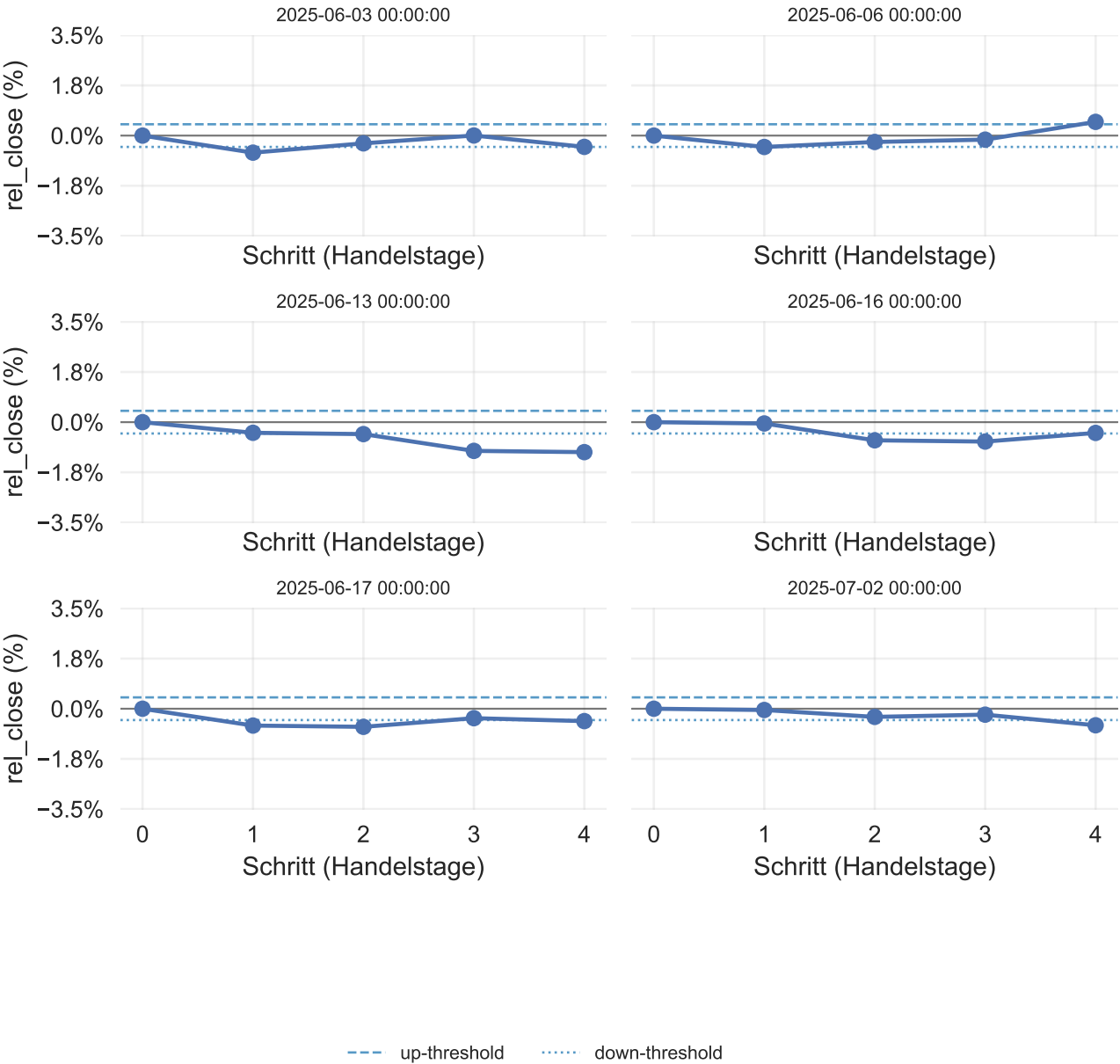


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

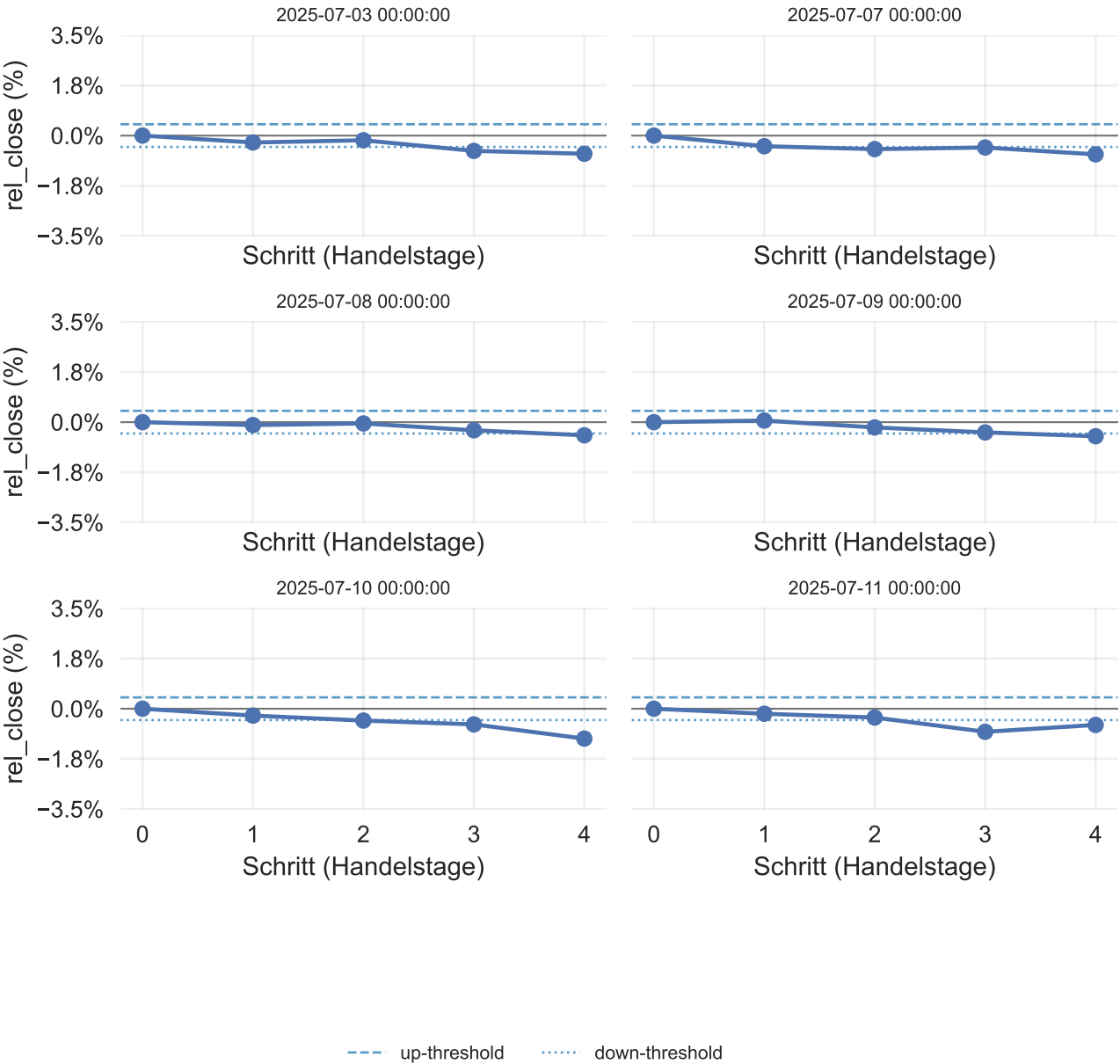


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

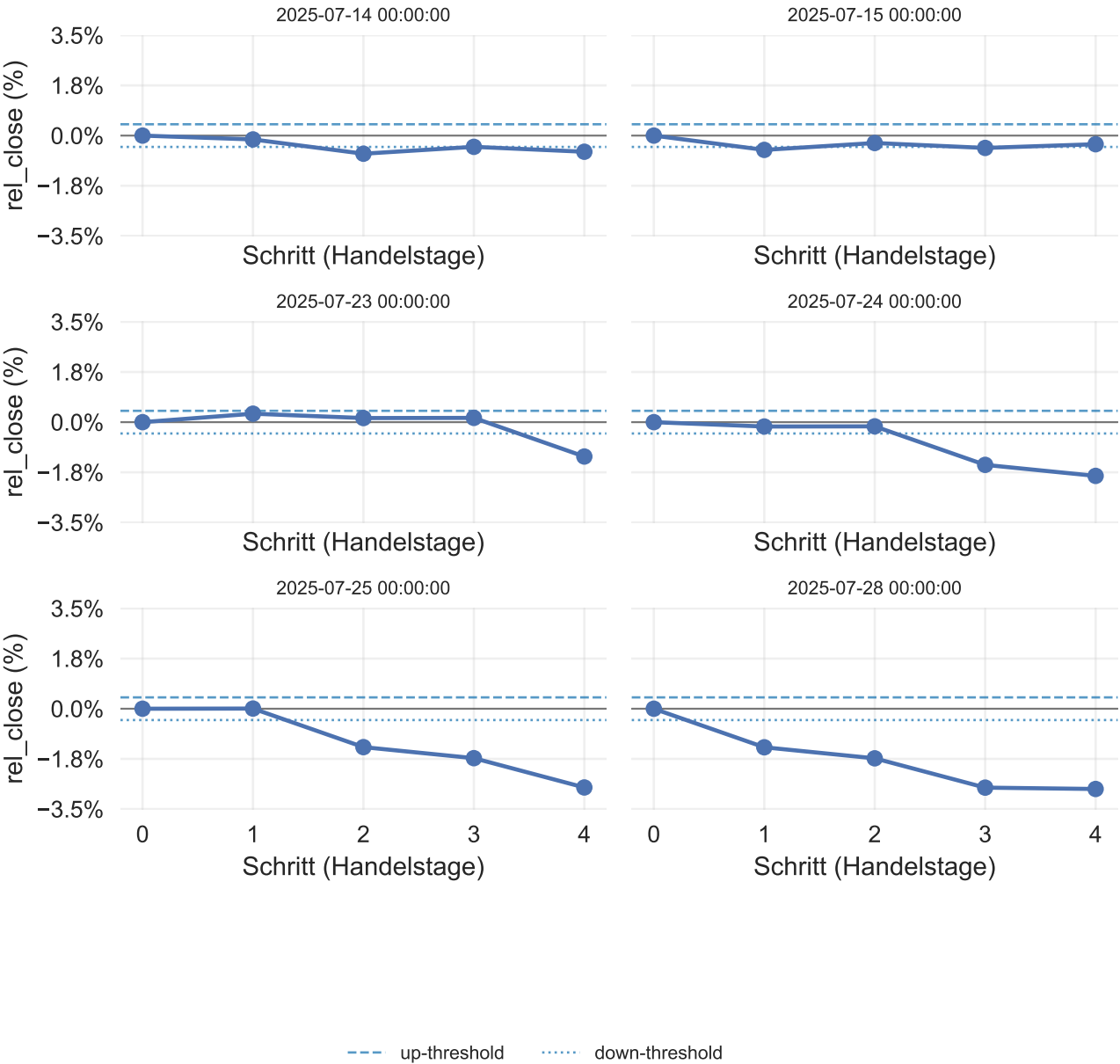


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

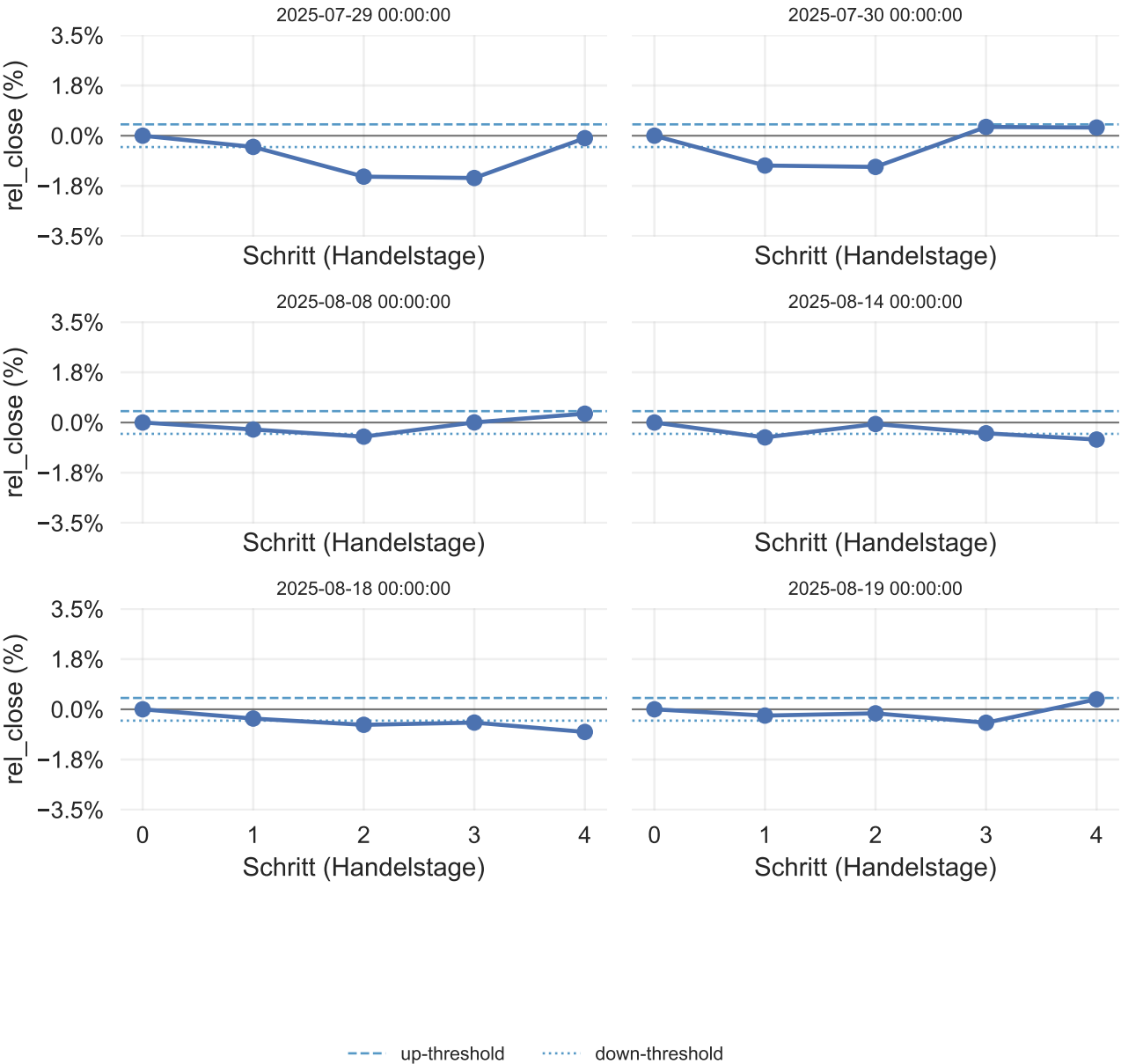


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.



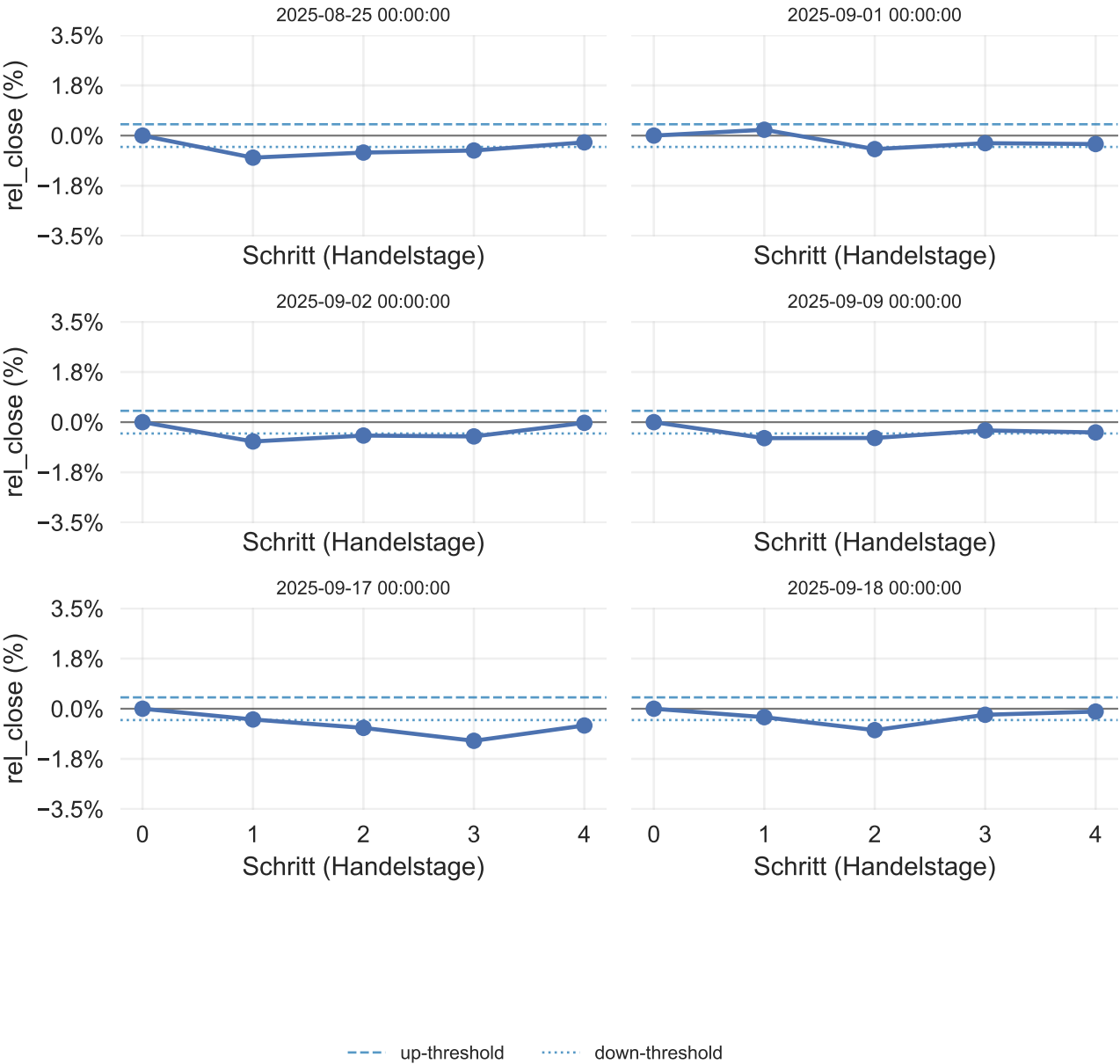


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

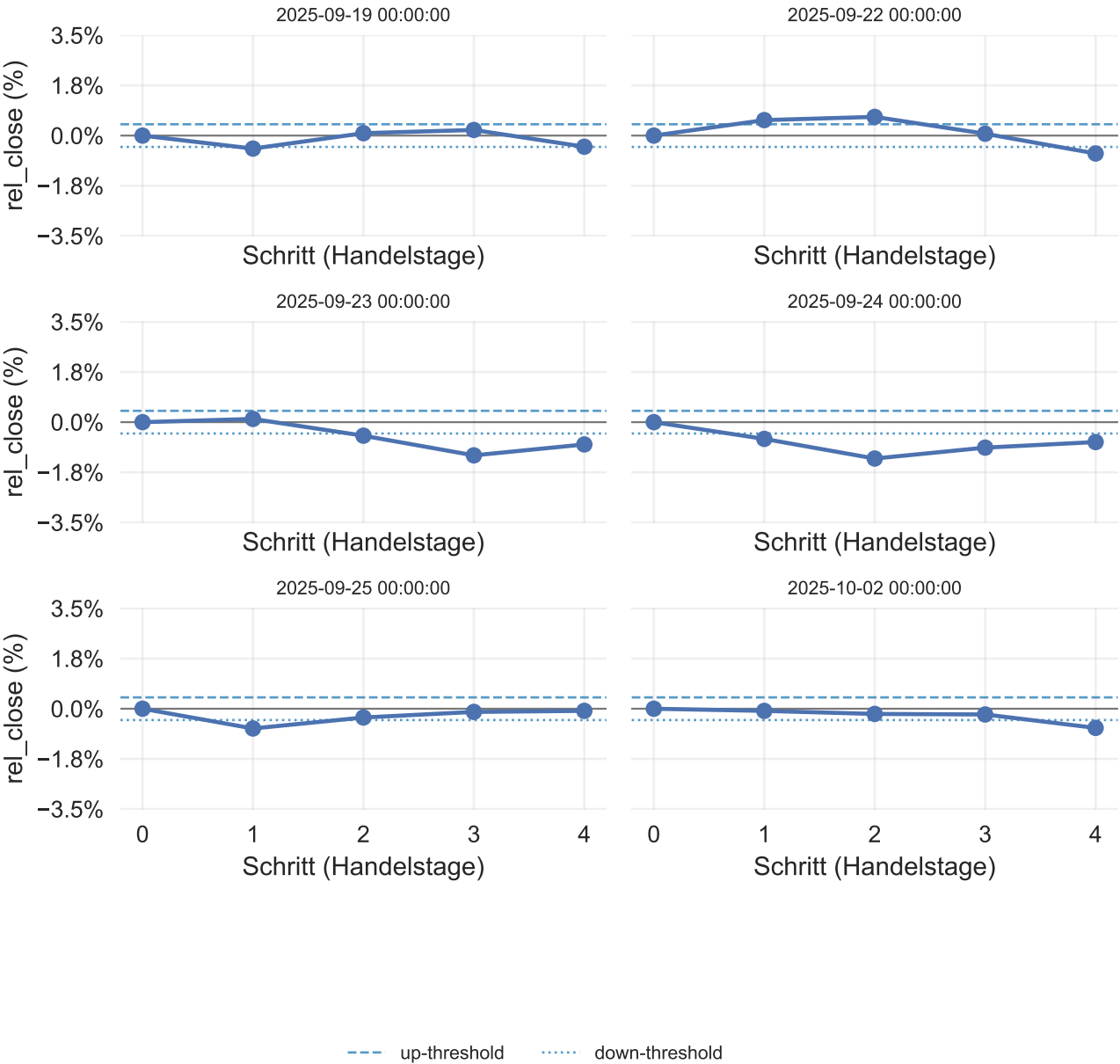


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

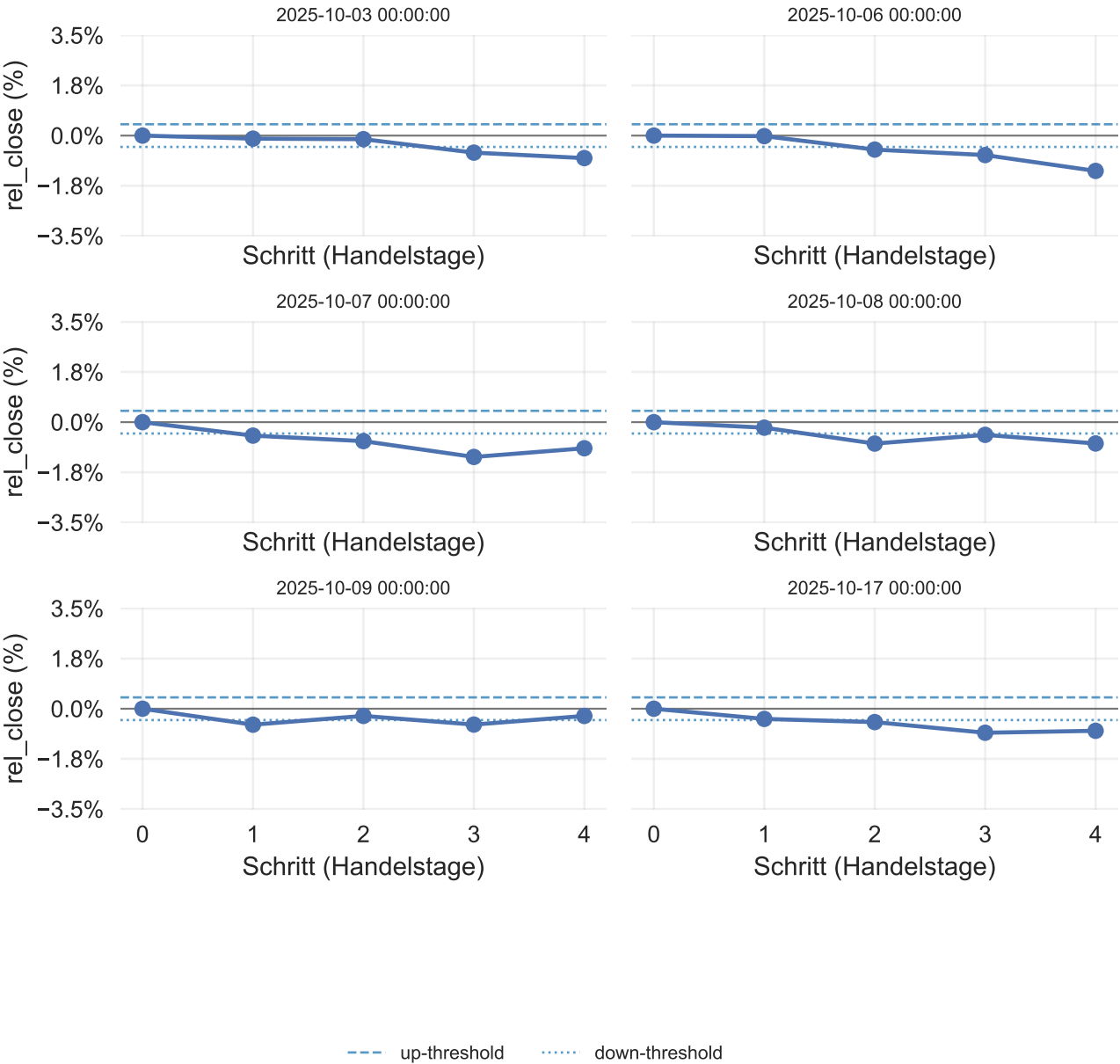


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

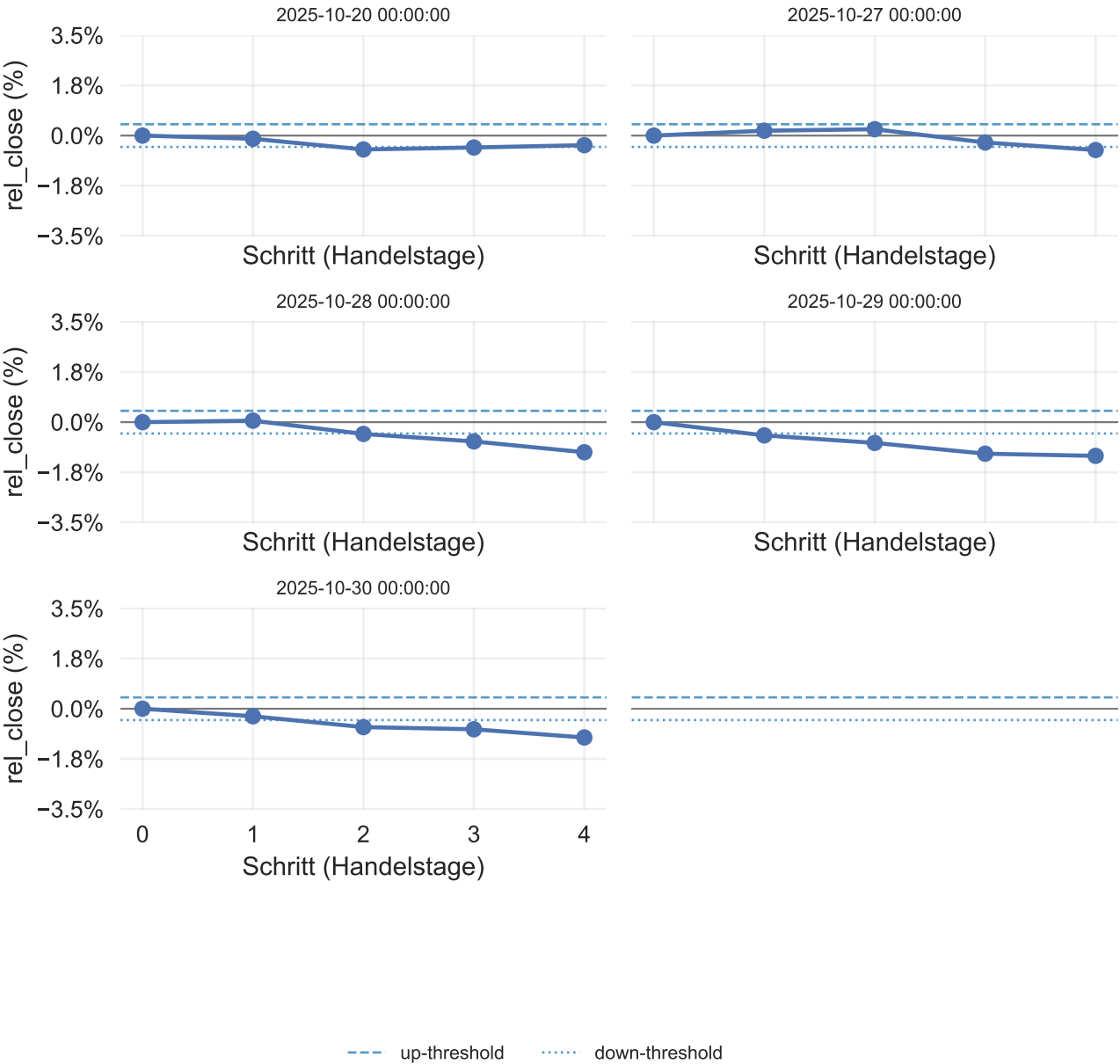


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

Signal-Modell – Kennzahlen für Klasse 'move' (train/val/test, thr=0.50)

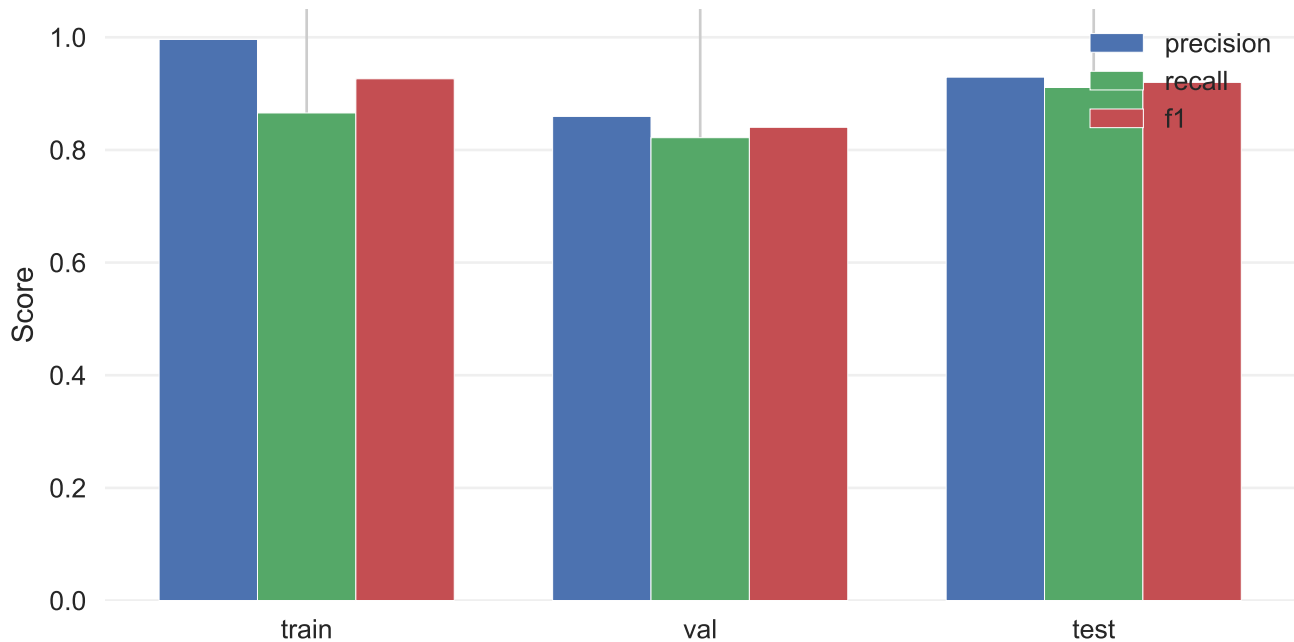


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

## Signal-Modell – Tabelle (Klasse 'move', thr=0.50)

split	precision	recall	f1	support
train	0.996	0.866	0.926	1809.000
val	0.860	0.822	0.840	432.000
test	0.929	0.911	0.920	202.000

Tabelle: Kennzahlen der positiven Klasse (precision/recall/F1/support) für train/val/test.

Richtungs-Modell – Kennzahlen für Klasse 'up' (train/val/test)

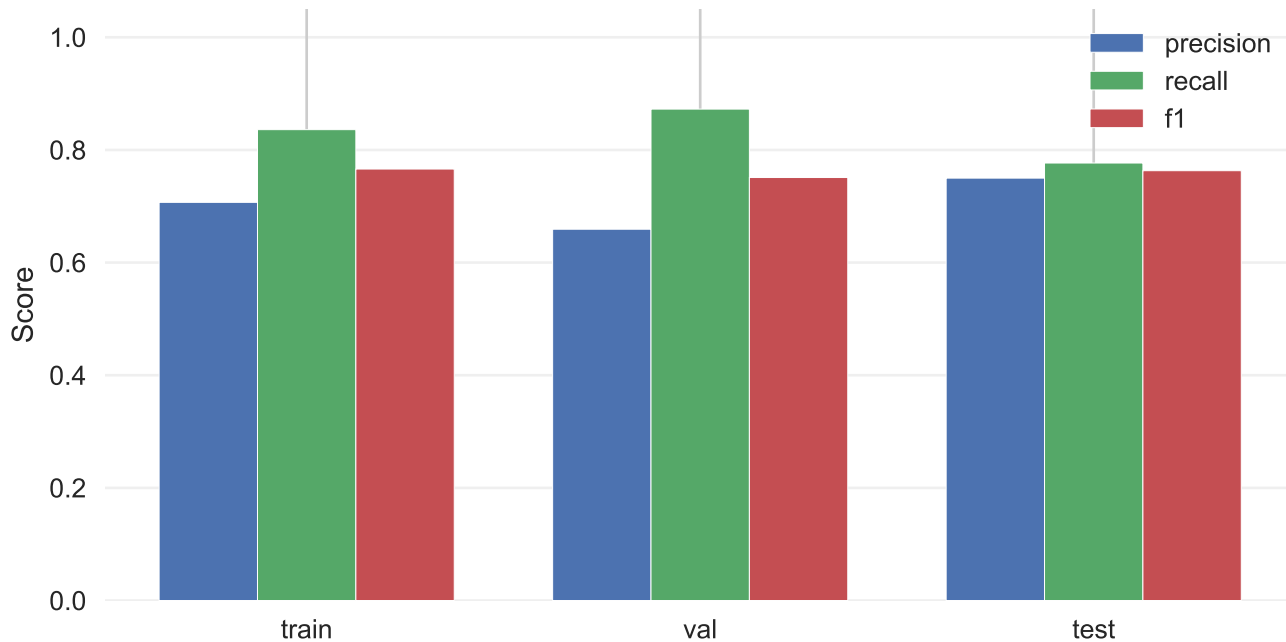


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

## Richtungs-Modell – Tabelle (Klasse 'up')

split	precision	recall	f1	support
train	0.707	0.836	0.766	860.000
val	0.659	0.873	0.751	204.000
test	0.750	0.777	0.763	112.000

Tabelle: Kennzahlen der positiven Klasse (precision/recall/F1/support) für train/val/test.



## Kombinierte Test-Auswertung – neutral / up / down

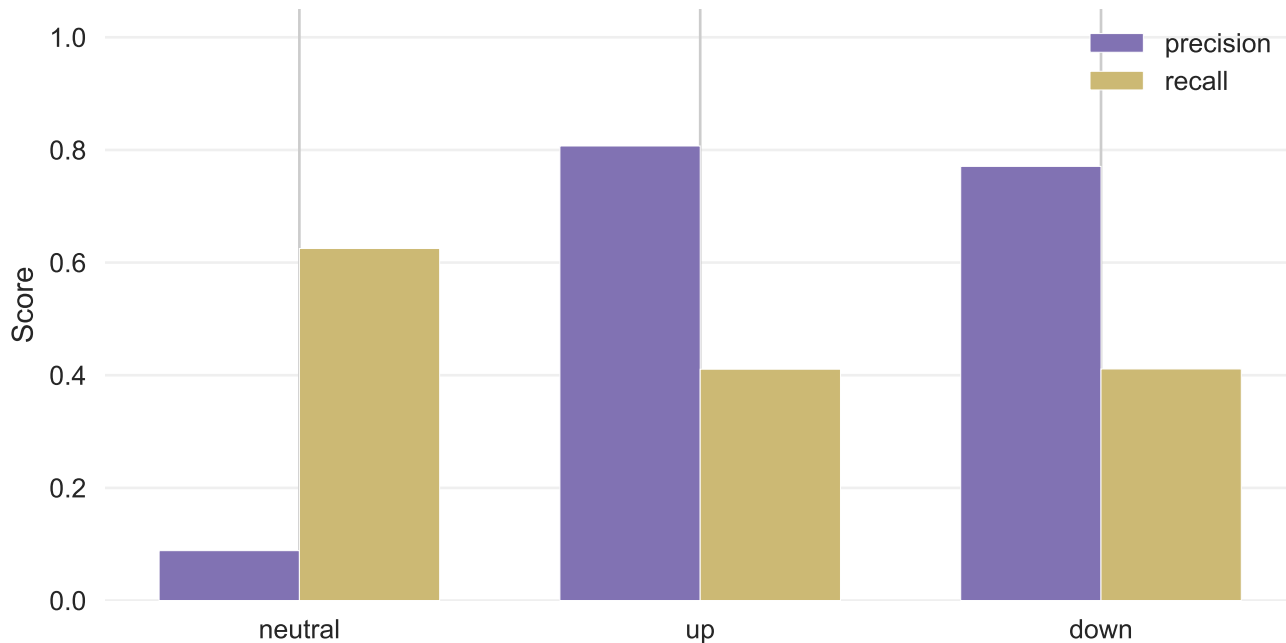


Abbildung: Precision und Recall der kombinierten 3-Klassen-Vorhersage (neutral/up/down) auf dem Test-Split.

## Kombiniertes Modell – Tabelle (Test, neutral/up/down)

klasse	precision	recall	f1	support
neutral	0.088	0.625	0.155	16
up	0.807	0.411	0.544	112
down	0.771	0.411	0.536	90

Tabelle: Kennzahlen der drei Klassen (neutral/up/down) des kombinierten Modells auf dem Test-Split.

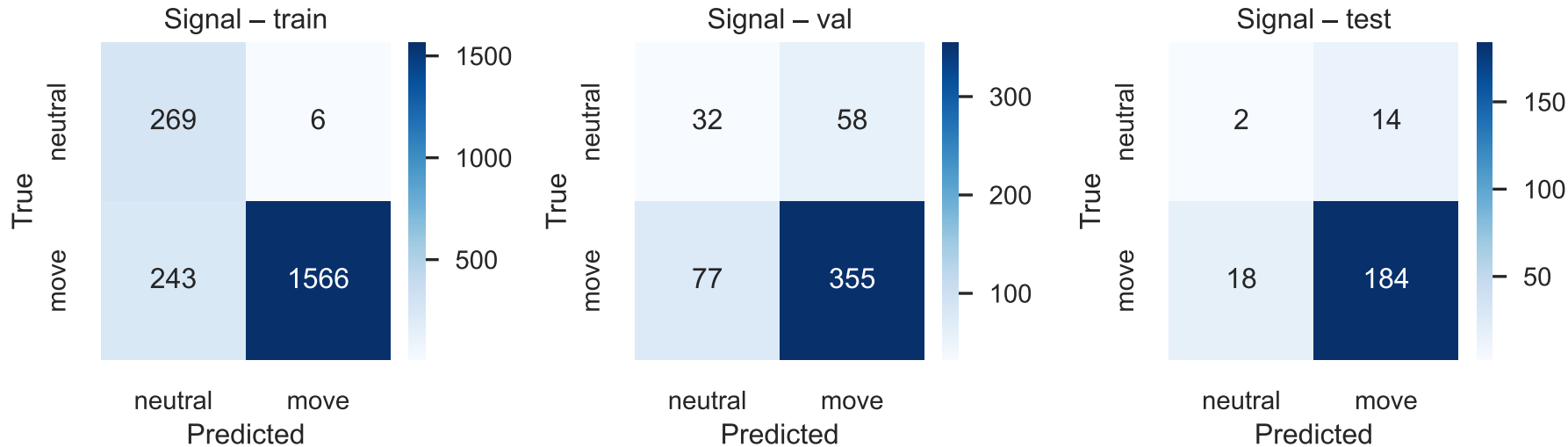


Abbildung: Confusion-Matrizen des Signal-Modells (neutral vs move) für Train-, Validierungs- und Test-Split.

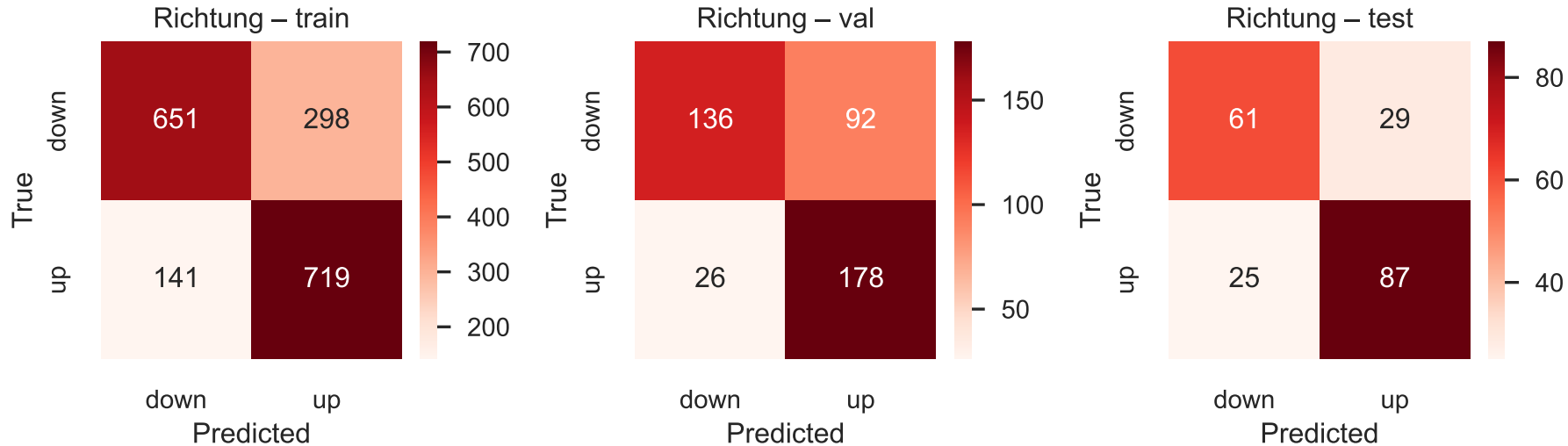


Abbildung: Confusion-Matrizen des Richtungs-Modells (down vs up) für Train-, Validierungs- und Test-Split.

Confusion Matrix – Test (neutral / up / down)

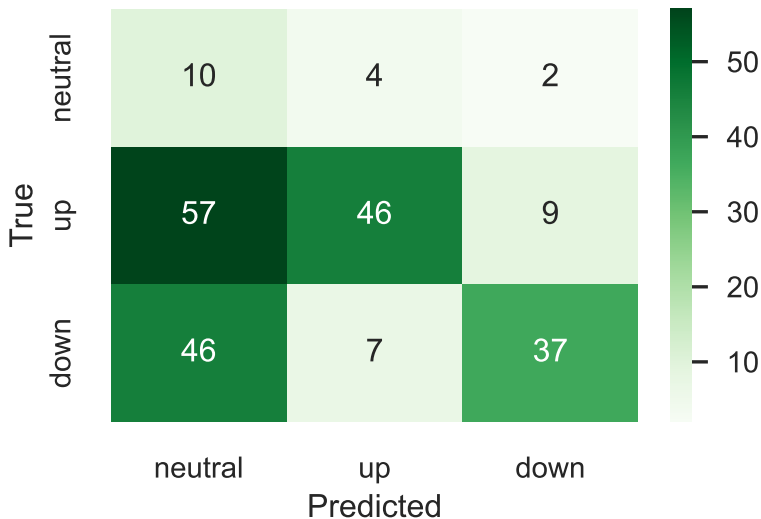


Abbildung: Confusion-Matrix des kombinierten Modells (neutral/up/down) auf dem Test-Split.

## Konfusionsmatrizen – Zählwerte (TN/FP/FN/TP)

modell	split	TN	FP	FN	TP
signal	train	269	6	243	1566
signal	val	32	58	77	355
signal	test	2	14	18	184
direction	train	651	298	141	719
direction	val	136	92	26	178
direction	test	61	29	25	87

Tabelle: Zählwerte der Konfusionsmatrizen (TN/FP/FN/TP)  
für Signal- und Richtungs-Modell je Split.

## Fehlklassifikationen – Übersicht (False Positives)

task	predicted	total_fp	true_label_breakdown
combined	up	11	neutral:4, down:7
combined	down	11	neutral:2, up:9
signal	move	14	neutral:14

Tabelle: Zusammenfassung der wichtigsten False-Positive-Fälle für kombinierten Test (neutral/up/down) und Signal-Test (neutral vs move).

Fehlklassifikationen (combined) im Test – Positionen auf der Preiszeitreihe (n=125/218 = 57.3%)

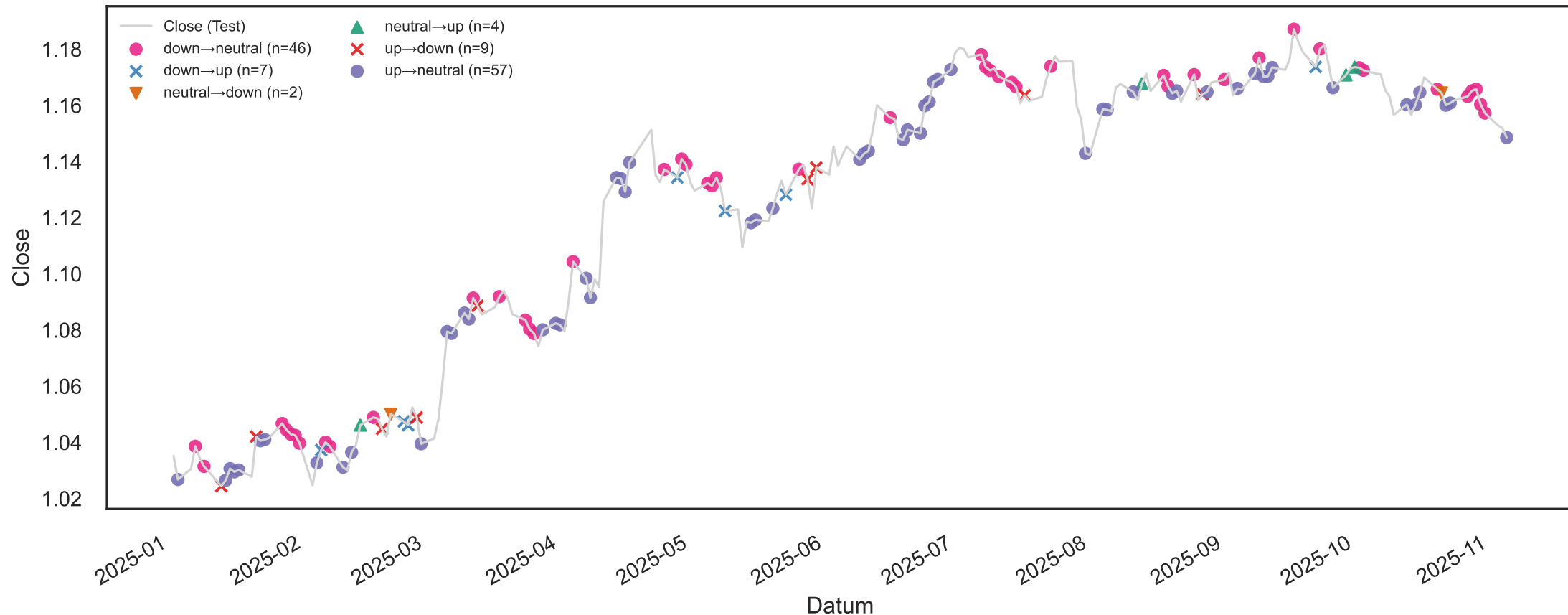


Abbildung: Jede Markierung ist ein Testtag, an dem der kombinierte Output (combined\_pred) vom true label abweicht. Farben/Marker zeigen den Fehlertyp true→pred.



## Signal-False-Positives im Test – Positionen auf der Preiszeitreihe

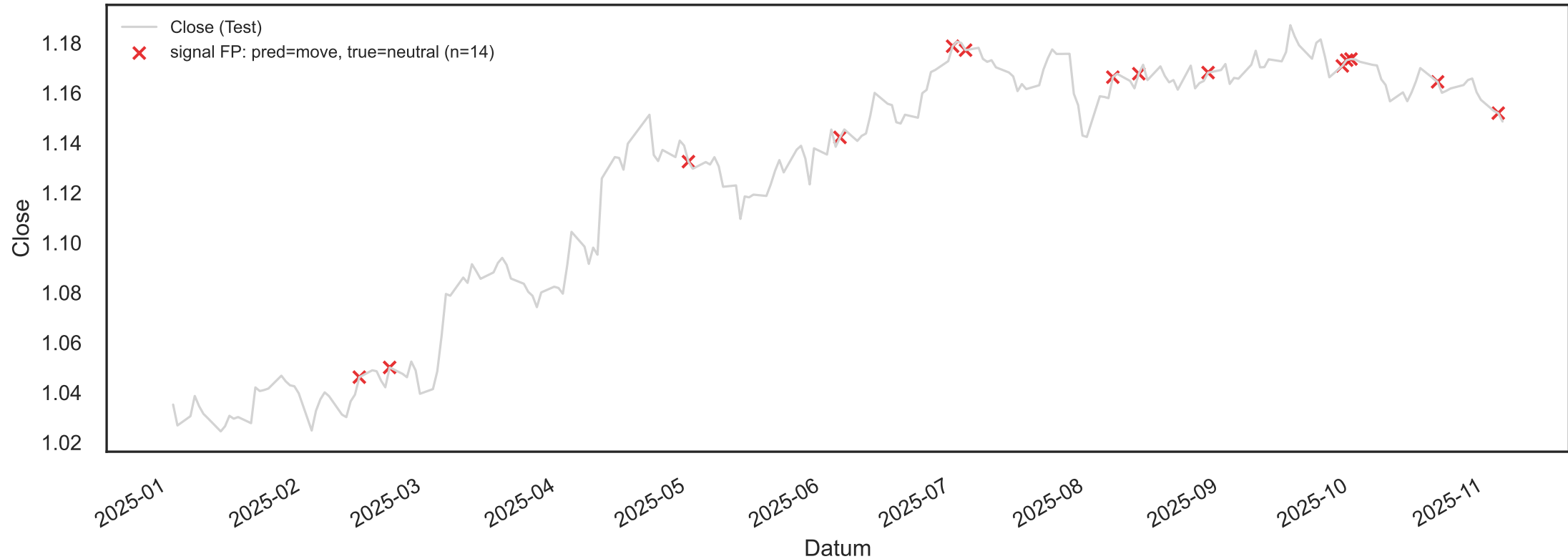


Abbildung: Markierte Testtage, an denen das Signal-Modell (neutral vs move) fälschlich ein Trade-Signal gegeben hat (pred=move), obwohl der Tag im Labeling neutral ist.

# Relativer Verlauf der Segmente (label='neutral→up')

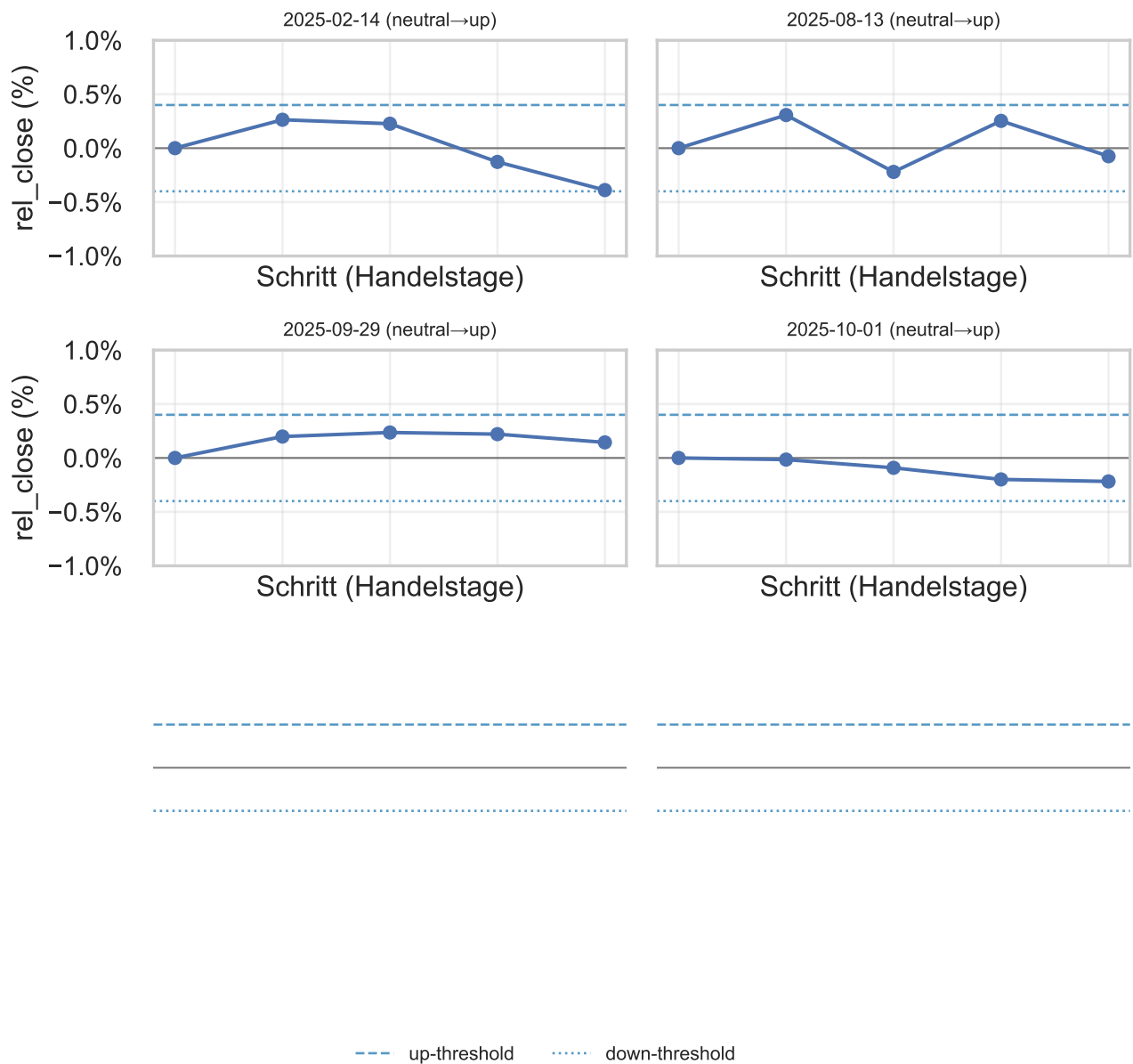


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'up' klassifiziert wurden.

Relativer Verlauf der Segmente (label='neutral→down')



Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'down' klassifiziert wurden.

## Tradesimulation – Strategien A und B (Test-Split)

Strategy	Kennzahl	Wert
A (fixer Einsatz)	Anzahl Trades	105
A (fixer Einsatz)	Einsatz up / down (CHF)	100 / 100
A (fixer Einsatz)	Trades up / down	57 / 48
A (fixer Einsatz)	Gewinner / Verlierer	90 / 15
A (fixer Einsatz)	Gesamt-P&L (CHF)	24.76
A (fixer Einsatz, Hebel 20)	Gesamt-P&L (CHF)	495.16
B (10% vom Kapital)	Startkapital (CHF)	1000.00
B (10% vom Kapital)	Endkapital (CHF)	1025.05
B (10% vom Kapital)	Minimum Kapital (CHF)	1000.40
B (10% vom Kapital, Hebel 20)	Endkapital (CHF)	1632.89
B (10% vom Kapital, Hebel 20)	Minimum Kapital (CHF)	1008.00

Tabelle: Zusammenfassung der Tradesimulation auf dem Test-Split.  
 Strategie A: fixer Einsatz pro Trade (100 CHF bei up, 100 CHF bei down).  
 Strategie B: 10 % des aktuellen Vermögens pro Trade (optional mit Hebel 20).

## Kostenmatrix – durchschnittliche Kosten pro Fall (Strategie A, Test-Split)

label_true	combined_pred	mean_chf
neutral	neutral	0.0
neutral	up	-1.0
neutral	down	-1.0
up	neutral	0.0
up	up	0.3999999999999998
up	down	-0.30608262914773204
down	neutral	0.0
down	up	0.04469813807531452
down	down	0.40000000000000001

Tabelle: durchschnittliche Kosten (CHF) pro Fall für jede Kombination aus wahren Label und vorhergesagtem Label (Strategie A, fixer Einsatz).

## Kostenmatrix – Gesamtkosten und Anzahl Trades (Strategie A, Test-Split)

label_true	combined_pred	count	sum_chf
neutral	neutral	10	0.0
neutral	up	4	-4.0
neutral	down	2	-2.0
up	neutral	57	0.0
up	up	46	18.399999999999999
up	down	9	-2.7547436623295884
down	neutral	46	0.0
down	up	7	0.3128869665272016
down	down	37	14.800000000000002

Tabelle: Anzahl Fälle und Gesamt-P&L (CHF) auf dem Test-Split  
für jede Kombination aus wahrem Label und vorhergesagtem Label (Strategie A).

## Strategie A vs B – Verlauf des Kapitals (ohne Hebel, Test-Split)

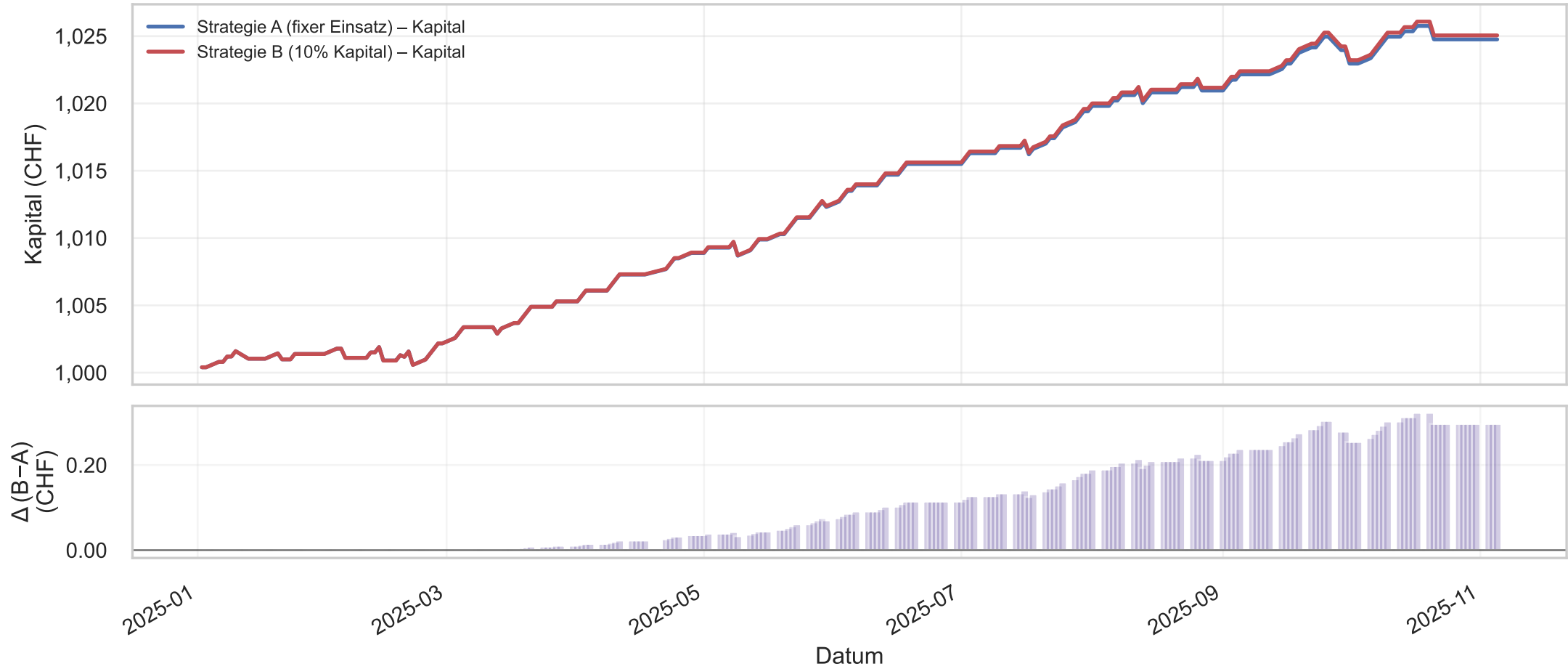


Abbildung: Oben Kapitalverlauf (CHF) für Strategie A und B ohne Hebel. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.

Strategie A vs B – kumulierter P&L (ohne Hebel, Test-Split)

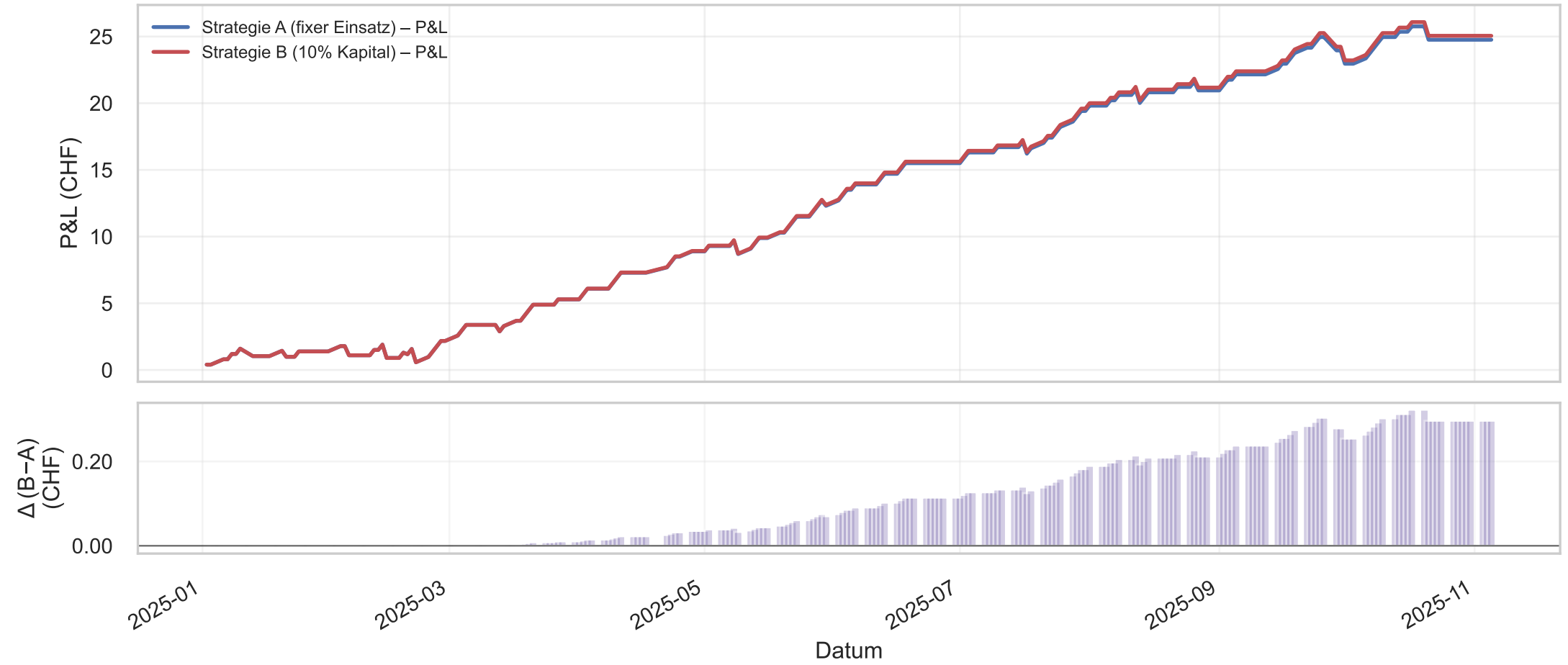


Abbildung: Oben kumulierter Gewinn/Verlust (P&L, CHF) für Strategie A und B ohne Hebel. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.



## Strategie A vs B – kumulierter Gewinn (P&L) als Punkte (ohne Hebel, Test-Split)

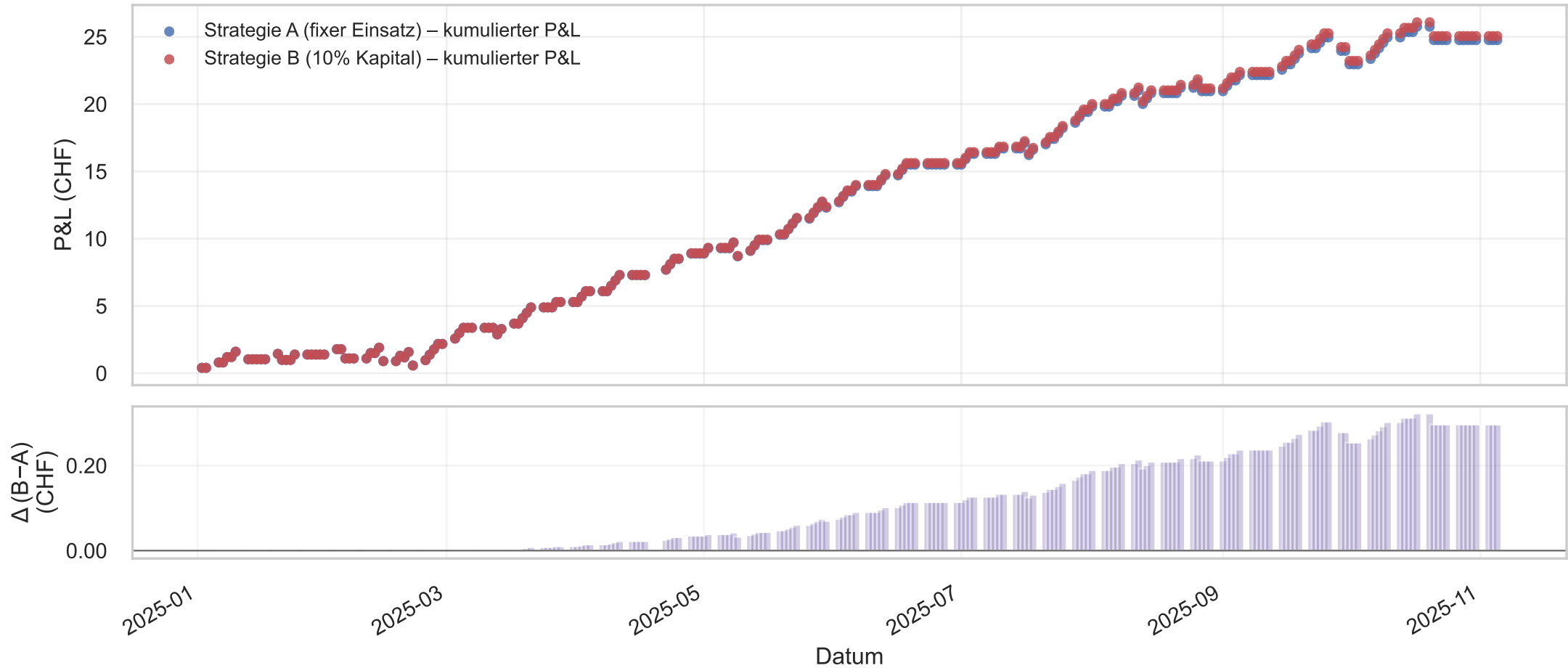


Abbildung: Oben kumulierter Gewinn/Verlust (P&L) als Punkte. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.

## Strategie A vs B – Verlauf des Kapitals (Hebel 20, Test-Split)

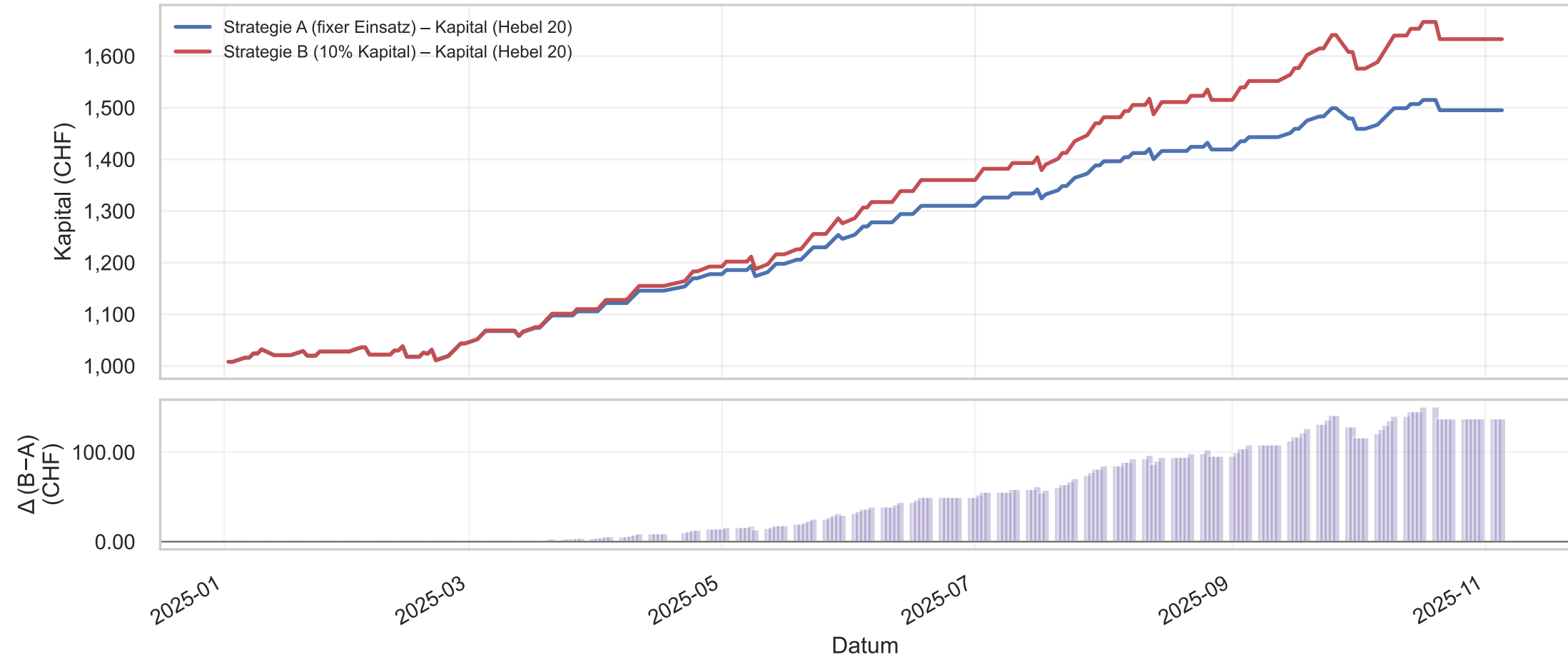


Abbildung: Oben Kapitalverlauf (CHF) für Strategie A und B mit Hebel 20. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.

## Strategie A vs B – kumulierter P&L (Hebel 20, Test-Split)

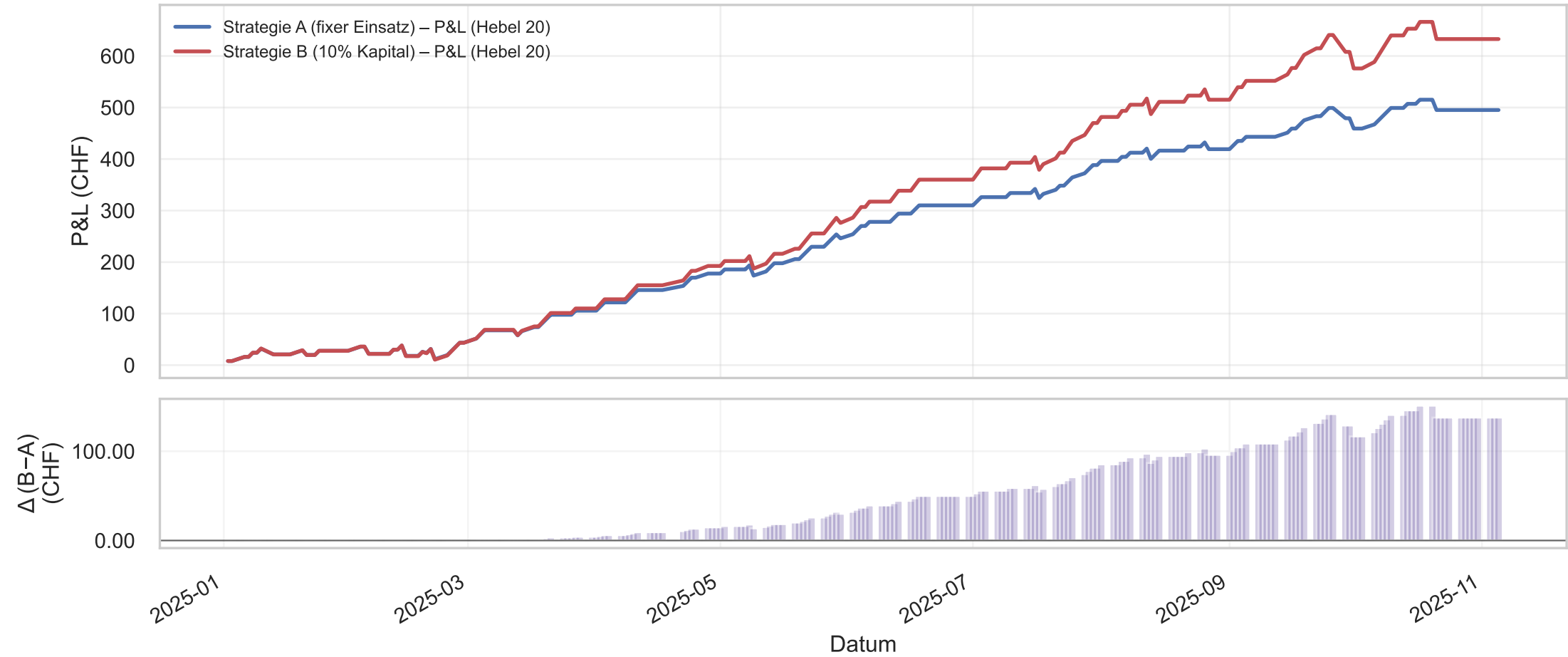


Abbildung: Oben kumulierter Gewinn/Verlust (P&L, CHF) für Strategie A und B mit Hebel 20. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.

## Strategie A vs B – kumulierter Gewinn (P&L) als Punkte (Hebel 20, Test-Split)

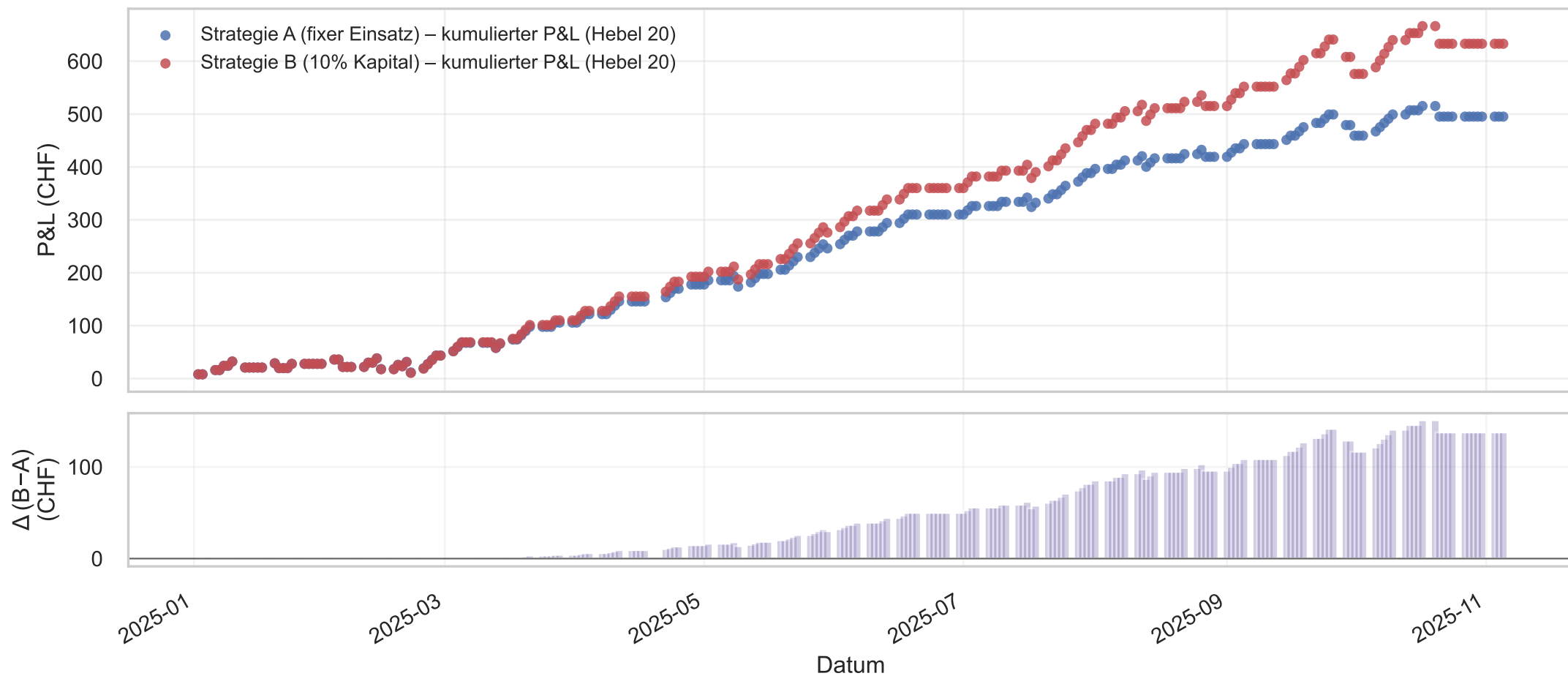
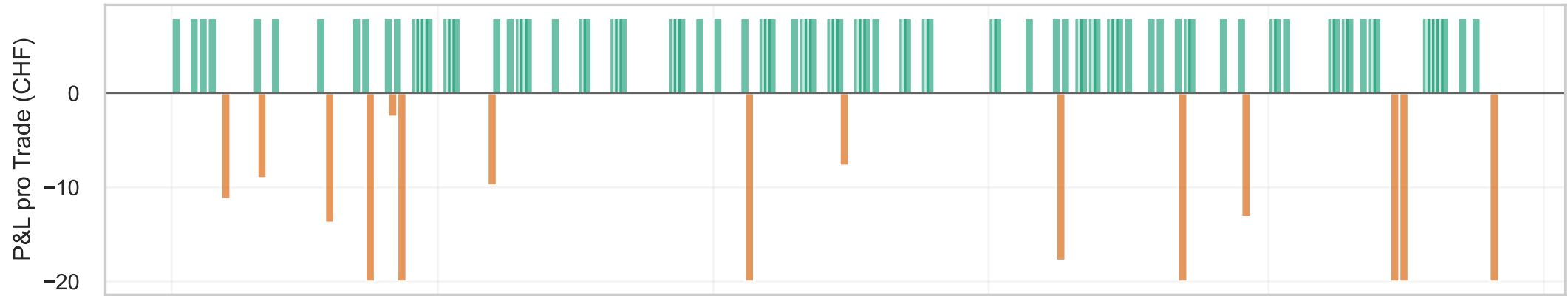
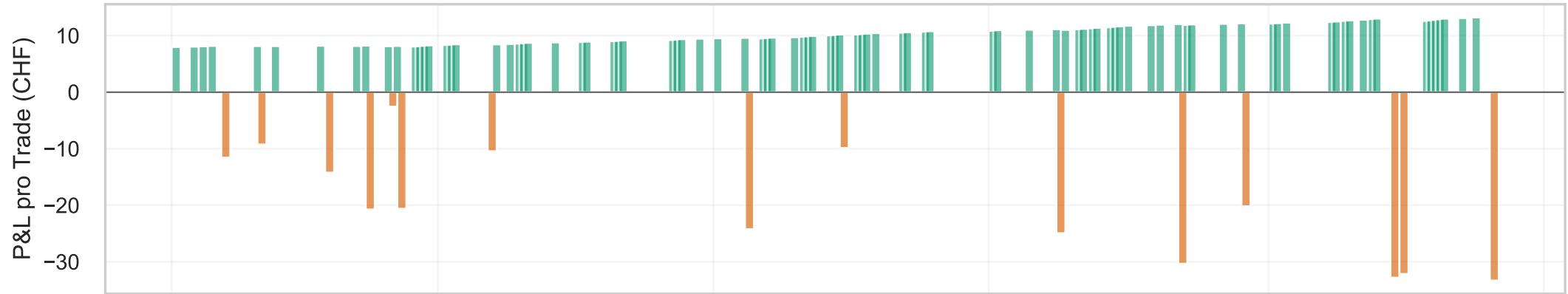


Abbildung: Oben kumulierter Gewinn/Verlust (P&L) als Punkte. Unten Balken: Differenz  $\Delta = (B - A)$  je Tag.

Strategie A – Gewinn pro Trade (Hebel 20, nur Trade-Tage)



Strategie B – Gewinn pro Trade (Hebel 20, nur Trade-Tage)



2025-01

2025-03

2025-05

2025-07

2025-09

2025-11

Datum

### Gewinn pro Monat (Hebel 20, Test-Split)

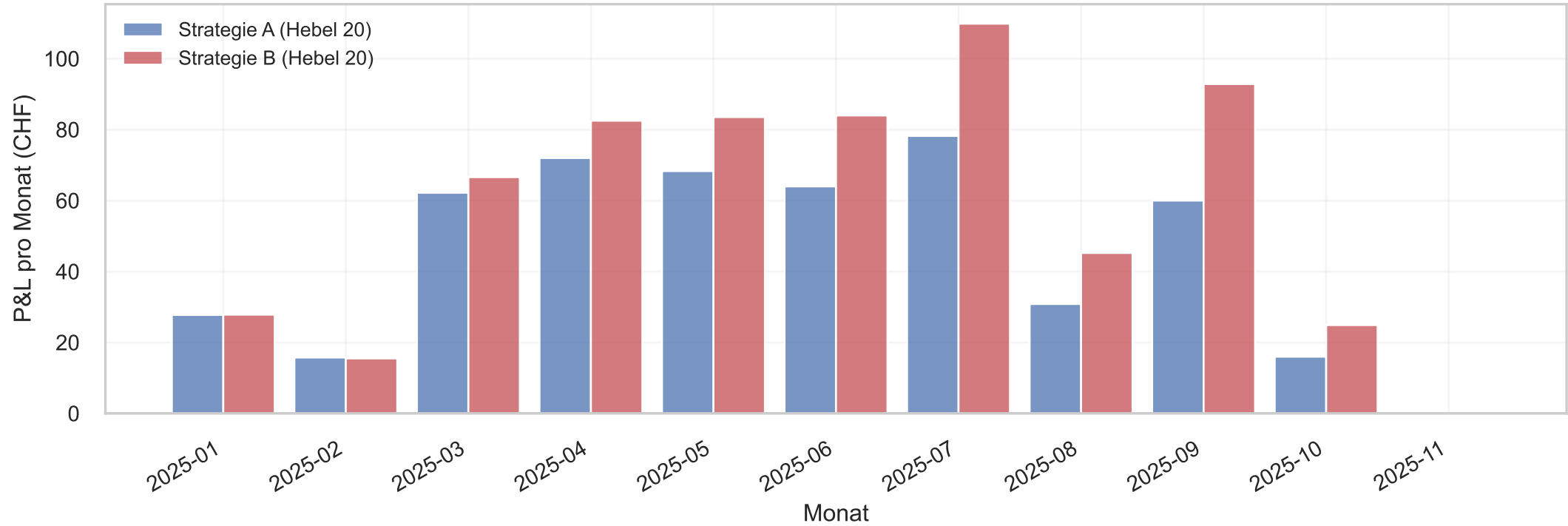


Abbildung: Summe der Tages-P&L je Monat. Hebel 20 ist bereits eingerechnet.

### 5-Jahres-Projektion (Bootstrap-Monte-Carlo, Hebel 20)

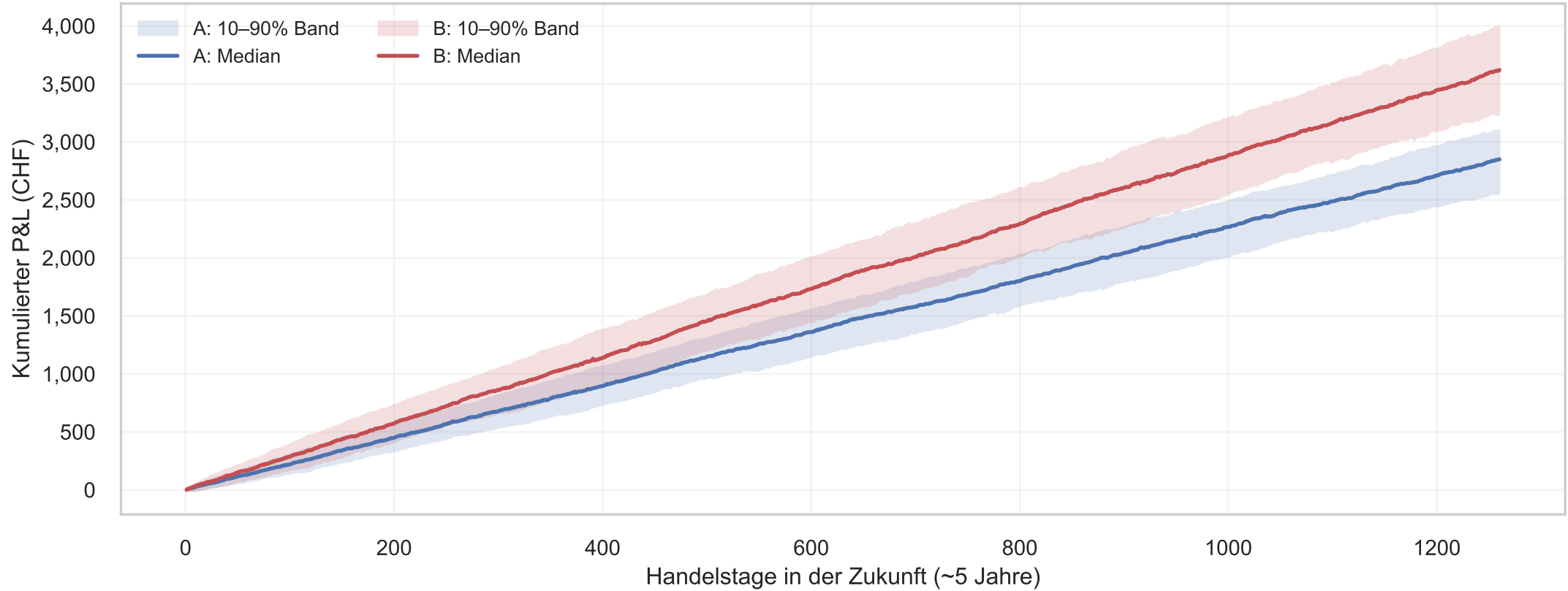


Abbildung: Keine echte Prognose. Es wird angenommen, dass die Verteilung der Tages-Ergebnisse aus dem Testzeitraum (inkl. Tage ohne Trades) in der Zukunft ähnlich bleibt. Gezeigt sind Median und 10-90%

## Strategie A – kumulierter P&L (Test-Split)

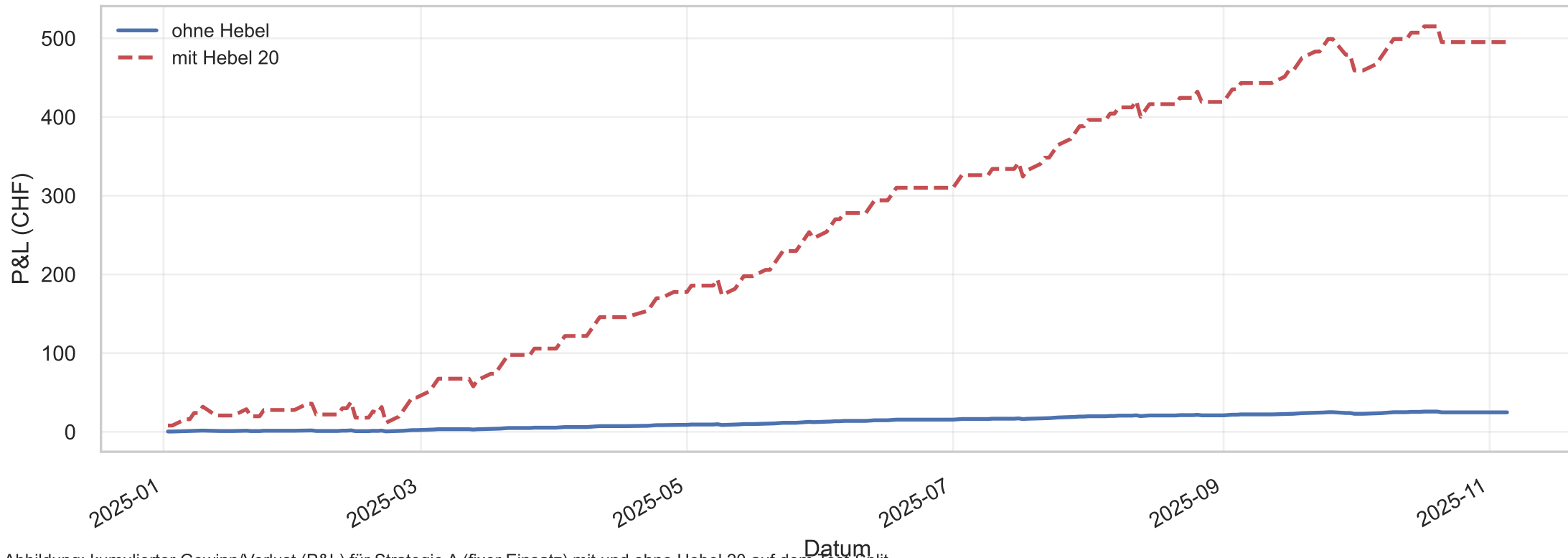


Abbildung: kumulierter Gewinn/Verlust (P&L) für Strategie A (fixer Einsatz) mit und ohne Hebel 20 auf dem Test-Split.



### Strategie B – kumulierter P&L (Test-Split)

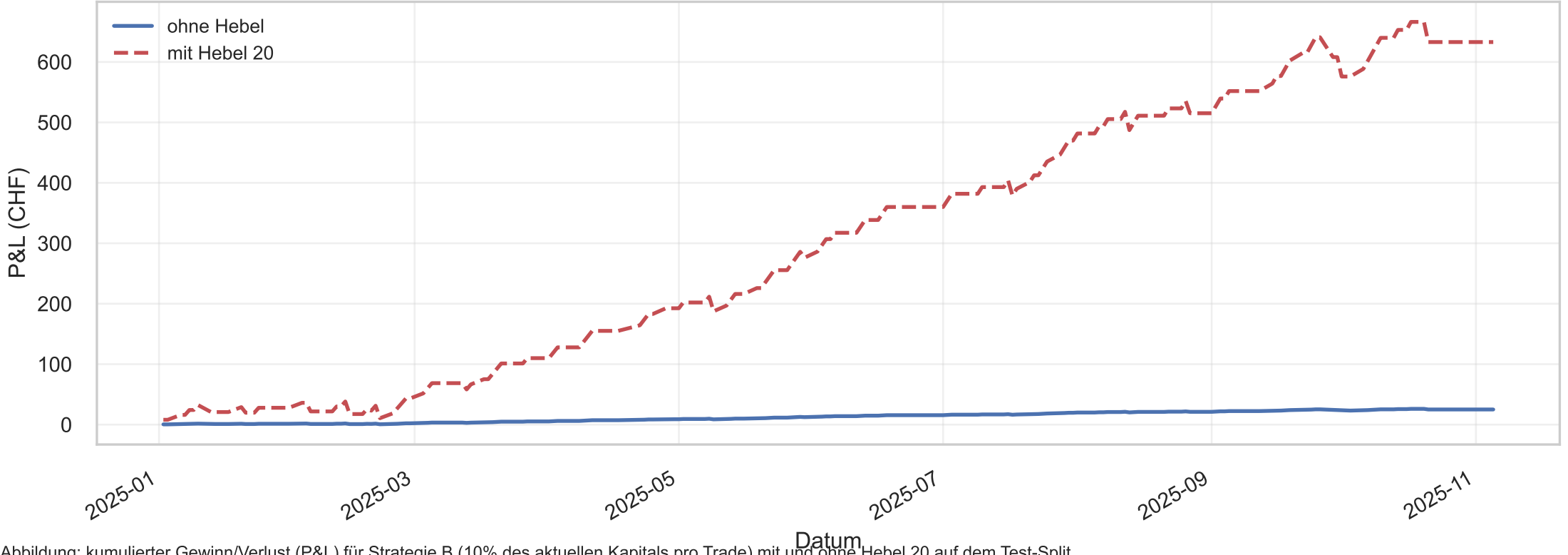


Abbildung: kumulierter Gewinn/Verlust (P&L) für Strategie B (10% des aktuellen Kapitals pro Trade) mit und ohne Hebel 20 auf dem Test-Split.

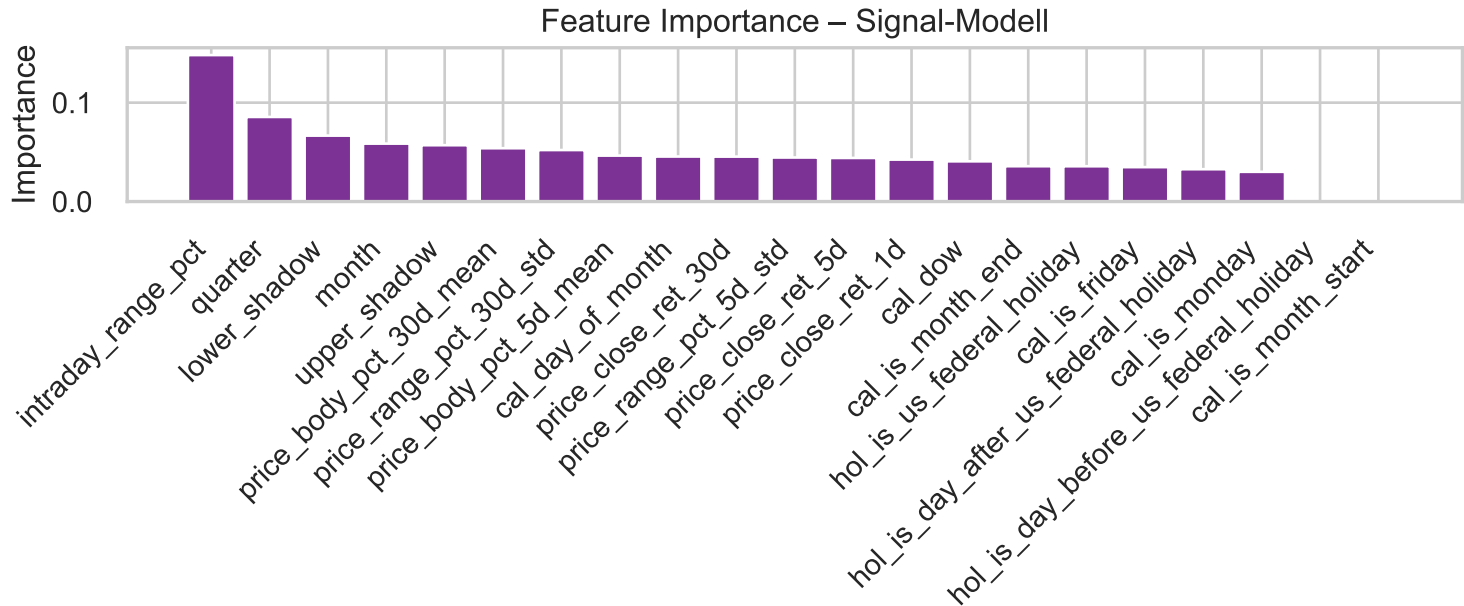


Abbildung: Wichtigkeit der Features für das Signal-Modell (neutral vs move).

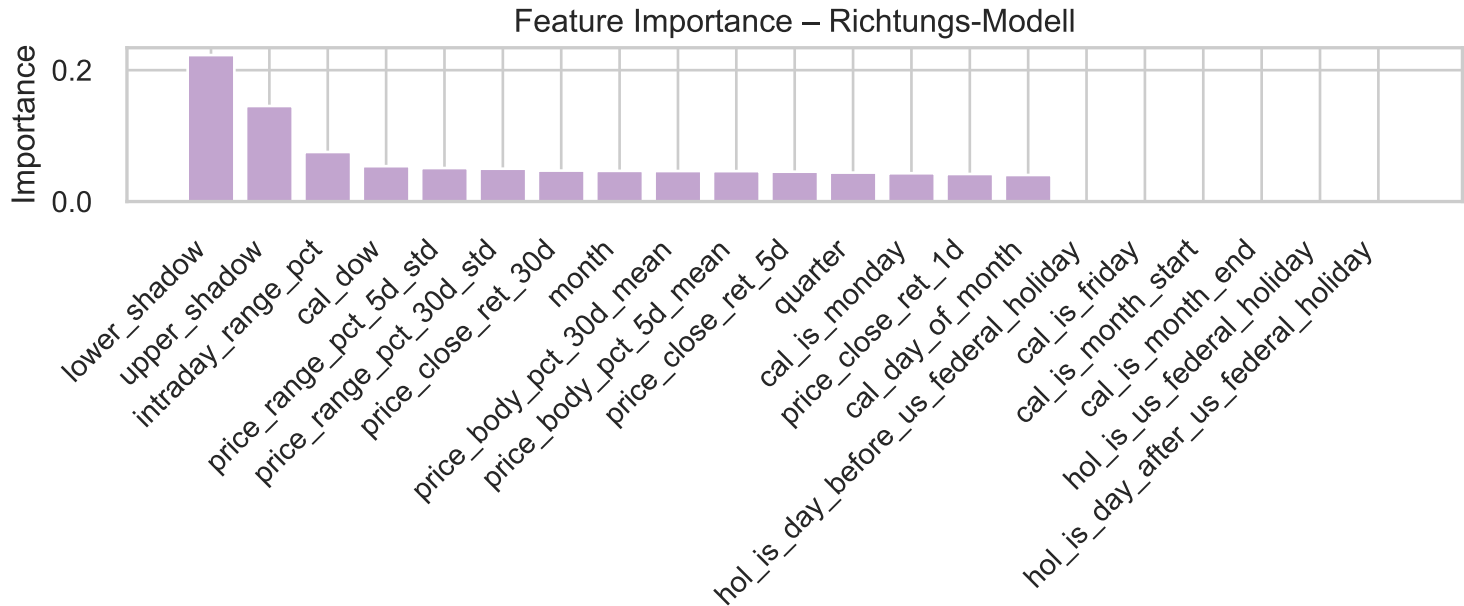


Abbildung: Wichtigkeit der Features für das Richtungs-Modell (down vs up).