Magic Numbers or Market Noise? Deconstructing the TD Setup in Time Series Predictions

Sofien Kaabar, CFA
sofien-kaabar@hotmail.com

December 27, 2025

Abstract

The TD Sequential indicator, specifically the "Setup" phase consisting of 9 consecutive closes with specific conditions every 4th bar, is a widely utilized counter-trend signal in technical analysis. However, the rigidity of the parameters ($n = 9$ setup, $k = 4$ lag) raises questions regarding parameter optimization and overfitting. This study investigates the predictive efficacy of the standard TD Setup (9, 4) against an alternative arbitrary pattern (8, 3) across a diversified basket of G10 currency pairs.

Using a standardized event study methodology, we isolate bullish and bearish reversal signals generated by both patterns. We evaluate performance using a number of metrics described in the research (e.g. statistical significance tests and bar efficiency). Our objective is to determine whether the canonical (9, 4) parameters possess unique predictive alpha or if a random variation (8, 3) yields statistically indistinguishable results, thereby challenging the robustness of the specific integer counts traditionally assigned to the pattern.

1 Introduction

Technical analysis often relies on specific integer constraints—such as the 14-period RSI or the 200-period Moving Average—that have become industry standards more through repetition than rigorous statistical validation. One of the most prominent examples of this is the **TD Sequential** indicator, developed by Tom DeMark. The core component of this system is the "Setup," which identifies potential trend exhaustion points. The canonical model demands a sequence of **9** consecutive candles meeting specific criteria.

Traders frequently debate whether these specific parameters represent a "golden ratio" of market psychology or simply a fitted artifact and an illusion. If the number 9 is truly significant, the pattern should significantly outperform a slightly looser or tighter constraint. If it is not, a pattern based on **8** consecutive candles should perform similarly or better due to the increased sample size of signals.

This report compares the predictive ability of the standard **TD Setup (9, 4)** against a modified **Setup (8, 3)**. By applying these rules to a basket of diversified currency pairs, we aim to isolate the marginal utility of waiting for that "9th" candle. The results have direct implications for algorithmic execution: if the (8, 3) pattern performs equally well, traders waiting for the perfect (9, 4) setup may be incurring unnecessary opportunity costs.

2 Methodology

2.1 The Canonical TD Setup (9, 4)

The TD Sequential is a bi-directional indicator designed to anticipate trend exhaustions. For the purpose of this comparative study, we evaluate the **Imperfect Setup** (also known as the Unperfected Setup). We deliberately exclude the secondary "Intersection" or "Validation" rules to isolate the predictive efficacy of the counting mechanism itself.

2.1.1 Bullish Setup (Buy Signal)

The **Imperfect Bullish Setup** identifies potential exhaustion in a downtrend based strictly on price persistence. It is defined by a single condition:

1. **The Setup Phase (The "9"):** There must be nine consecutive price bars where the *Close* is lower than the *Close* four bars earlier.

$$Close_t < Close_{t-4} \quad \forall t \in \{1, \ldots, 9\} \tag{1}$$

The signal is triggered immediately upon the close of the 9th bar. If the condition is broken at any point (e.g., at bar 7), the count resets to zero.

2.1.2 Bearish Setup (Sell Signal)

The **Imperfect Bearish Setup** identifies potential exhaustion in an uptrend and is the inverse of the bullish variant:

1. **The Setup Phase (The "9"):** There must be nine consecutive price bars where the *Close* is higher than the *Close* four bars earlier.

$$Close_t > Close_{t-4} \quad \forall t \in \{1, \ldots, 9\} \tag{2}$$

The signal is triggered immediately upon the close of the 9th bar. A break in the sequence resets the count.

2.2 The Alternative Setup (8, 3)

To test the robustness of the "9" parameter, we define a challenger pattern: the **Alternative Setup**. As with the canonical model, we evaluate the imperfect variation without secondary validation filters.

2.2.1 Bullish Setup (Buy Signal)

The **Bullish Setup** identifies potential exhaustion in a downtrend based strictly on price persistence. It is defined by a single condition:

1. **The Setup Phase (The "8"):** There must be eight consecutive price bars where the *Close* is lower than the *Close* three bars earlier.

$$Close_t < Close_{t-3} \quad \forall t \in \{1, \ldots, 8\} \tag{3}$$

The signal is triggered immediately upon the close of the 8th bar. If the condition is broken at any point (e.g., at bar 5), the count resets to zero.

2.2.2 Bearish Setup (Sell Signal)

The **Bearish Setup** identifies potential exhaustion in an uptrend and is the inverse of the bullish variant:

1. **The Setup Phase (The "8"):** There must be eight consecutive price bars where the *Close* is higher than the *Close* three bars earlier.

$$Close_t > Close_{t-3} \quad \forall t \in \{1, \ldots, 8\} \tag{4}$$

The signal is triggered immediately upon the close of the 8th bar. A break in the sequence resets the count.

2.3 Evaluation Framework

To ensure statistical robustness and mitigate the risk of regime-specific bias, we employ a diversified basket of uncorrelated G10 currency pairs. The dataset comprises approximately 11,000 data points (hourly bars) for each pair. This extensive sample size is calibrated to generate between 200 and 300 independent signal events per currency pair for each pattern variation ($N_{signals} \approx 200 - 300$), providing sufficient statistical power for hypothesis testing.

We align all signals at $t = 0$ (the completion of the setup candle) and calculate the forward returns over three distinct time horizons to capture different reversal dynamics:

- **Short-Term Reversal ($k = 5$):** Captures the immediate reaction and potential "bounce" effect.

- **Intermediate Validation ($k = 10$):** Assesses the sustainability of the reversal.
- **Medium-Term Trend Change ($k = 20$):** Determines if the signal marked a significant local inflection point.

For each horizon, we compute a suite of metrics, including Profit Factor, Hit Ratio, Expectancy, and Bar Efficiency. Finally, we establish the statistical significance of the performance delta between the (9, 4) and (8, 3) setups using the Two-Sample t-Test for mean returns and the Kolmogorov-Smirnov (K-S) Test for distributional equivalence.

3 Performance Evaluation Metrics

To rigorously benchmark the canonical TD Setup against its accelerated counterpart, we move beyond simple profitability measures. The following metrics were selected to quantify not just the outcome of the trades, but the quality, efficiency, and statistical validity of the signals.

3.1 Hit Ratio (Win Rate)

The Hit Ratio measures the directional accuracy of the signal. It is defined as the percentage of signals that result in a positive return ($R_{t+k} > 0$) at the end of the holding period.

$$\text{Hit Ratio} = \frac{N_{wins}}{N_{total}} \tag{5}$$

While a high win rate is desirable, it is not sufficient for profitability on its own; it must be analyzed in conjunction with the payout ratio (Expectancy).

3.2 Profit Factor

The Profit Factor quantifies the risk-reward relationship of the strategy. It is calculated as the ratio of gross profits to gross losses.

$$\text{Profit Factor} = \frac{\sum \text{Gross Profit}}{\sum |\text{Gross Loss}|} \tag{6}$$

A value greater than 1.0 indicates a profitable system. In the context of this study, we compare the Profit Factors to determine if the 9-4 setup "filters" out losing trades more effectively than the 8-3 setup.

3.3 Expectancy

Expectancy represents the average amount a trader can expect to win (or lose) per trade, normalized by the frequency of wins and losses. It combines the Hit Ratio with the average size of wins and losses.

$$E = (P_{win} \times \text{Avg Win}) - (P_{loss} \times \text{Avg Loss}) \tag{7}$$

This metric is crucial for determining if the canonical setup offers a higher "edge" per transaction compared to the accelerated setup.

3.4 Bar Efficiency

Bar Efficiency measures the "return on time" invested in the trade. Since the 9-4 setup requires waiting for an additional candle compared to the 8-3 setup, it incurs a time cost. Bar Efficiency determines if the returns justify this wait.

$$\text{Efficiency} = \frac{\text{Net Profit}}{\text{Total Bars Held}} \tag{8}$$

If the 8-3 setup has a higher Bar Efficiency, it implies that waiting for the 9th candle offers diminishing marginal returns.

3.5 Two-Sample t-Test

To determine if the difference in profitability is real or a result of random noise, we employ the Welch's t-test. This statistical test compares the mean returns of the two independent groups (9-4 vs. 8-3).

- **Null Hypothesis (H_0):** $\mu_{9-4} = \mu_{8-3}$ (The average returns are identical).
- **Result:** A p-value < 0.05 indicates a statistically significant difference in performance.

3.6 Kolmogorov-Smirnov (K-S) Test

While the t-test compares averages, the K-S test compares the *shape* of the return distributions. It assesses whether the risk profile (volatility, tail risk, skewness) of the two setups differs significantly.

- **Null Hypothesis (H_0):** The two samples are drawn from the same distribution.
- **Result:** If the p-value is high (> 0.05), it confirms that the market behavior following an "8" count is statistically indistinguishable from the behavior following a "9" count.

4 Results

We applied the evaluation methodology across our diversified currency basket. The following section presents the visual confirmation of the signal generation process.

4.1 Signal Visualization

To validate the pattern recognition logic, we visualize specific signal instances. Figure 1 demonstrates the TD Setup (9-4) on USDCHF, where the indicator tries to identify local exhaustion points.

Figure 1: EURSEK Hourly Chart: 9-4 Pattern. This is the regular Tom DeMark's TD Sequential Pattern with the imperfect setup conditions. The signals are generated on the close prices upon generating the signal

Subsequently, Figure 2 illustrates the Alternative Setup (8-3) on EURSEK. The signals refer to expected intermediate turning points.

Figure 2: EURSEK Hourly Chart: 8-3 Pattern. Note the signal generation on the 8th count, allowing for an earlier entry compared to the canonical model.

Table 1: Comparative Performance Metrics for $k = 5$ (Short-Term Reversal)

	9-4 Pattern				**8-3 Pattern**				**Statistical Tests**	
Asset	**PF**	**Hit %**	**Exp.**	**Eff.**	**PF**	**Hit %**	**Exp.**	**Eff.**	**T-Test**	**KS-Test**
EURUSD	0.79	52.36%	-16.67	-3.33	0.94	48.66%	-3.79	-0.76	0.4745	0.5032
USDCHF	1.06	50.11%	3.96	0.79	1.02	53.29%	1.41	0.28	0.8854	0.6571
USDCAD	1.27	50.01%	12.27	2.45	1.05	50.47%	2.58	0.52	0.4615	0.9288
EURSEK	1.08	47.10%	4.40	0.88	0.95	50.42%	-2.98	-0.60	0.6307	0.7123
CHFJPY	0.80	46.82%	-15.99	-3.20	0.97	52.30%	-2.28	-0.46	0.4248	0.3711
EURHUF	1.38	57.14%	17.08	3.42	1.45	58.10%	21.53	4.31	0.7744	0.9036
USDMXN	0.99	50.36%	-1.53	-0.31	0.79	50.00%	-25.19	-5.04	0.4121	0.9802
USDSGD	0.69	49.83%	-16.59	-3.32	0.87	50.35%	-6.80	-1.36	0.3801	0.8013

Note: PF = Profit Factor; Exp. = Expectancy; Eff. = Bar Efficiency. All statistical tests yield p-values > 0.05, indicating no statistically significant difference between the two patterns.

4.2 Statistical Performance ($k = 5$)

Table 1 summarizes the performance metrics for a 5-day holding period. The data compares the Profit Factor, Hit Ratio, Expectancy, and Bar Efficiency for both the Canonical (9) and Accelerated (8) patterns. The final two columns present the p-values for the Two-Sample t-Test and the Kolmogorov-Smirnov (KS) Test.

The results indicate a high degree of convergence between the two patterns. Across all assets, the T-Test p-values range from 0.38 to 0.98, consistently failing to reject the null hypothesis. This suggests that the mean returns generated by the 9-count setup are statistically indistinguishable from those of the 8-count setup. Similarly, the KS-Test confirms that the risk distribution profiles are identical, challenging the necessity of the stricter 9-candle constraint. The performance metrics (e.g. Profit Factor and Expectancy) seem to be random between the two patterns.

4.3 Statistical Performance ($k = 10$)

Extending the holding period to 10 days allows us to evaluate the intermediate validation of the reversal signal. Table 2 presents the comparative metrics for this horizon.

Table 2: Comparative Performance Metrics for $k = 10$ (Intermediate Validation)

	9-4 Pattern				8-3 Pattern				Statistical Tests	
Asset	**PF**	**Hit %**	**Exp.**	**Eff.**	**PF**	**Hit %**	**Exp.**	**Eff.**	**T-Test**	**KS-Test**
EURUSD	0.86	50.51%	-9.56	-1.37	0.80	50.35%	-13.70	-1.37	0.7699	0.9778
USDCHF	0.92	49.68%	-9.58	-0.96	0.92	52.00%	-10.55	-1.05	0.9732	0.9305
USDCAD	1.10	49.81%	7.28	0.73	1.01	49.62%	0.46	0.05	0.7286	0.9565
EURSEK	1.23	53.28%	16.72	1.37	1.10	49.16%	7.34	0.73	0.6227	0.8814
CHFJPY	0.75	47.83%	-31.23	-3.12	1.00	50.53%	0.56	0.06	0.2059	0.1319
EURHUF	2.21	65.74%	51.98	5.20	2.04	59.81%	51.80	5.18	0.9927	0.5517
USDMXN	1.08	52.14%	11.17	1.12	1.01	48.51%	1.18	0.12	0.7902	0.6662
USDSGD	0.86	50.51%	-9.56	-0.96	0.80	50.35%	-13.70	-1.37	0.7699	0.9778

Note: Even at the intermediate horizon, T-Test values remain high. Notably, EURHUF shows strong performance on both patterns (PF > 2.0), yet the difference between the 9 and 8 counts remains statistically insignificant ($p = 0.99$).

4.4 Statistical Performance ($k = 20$)

Finally, Table 3 examines the medium-term trend change potential over a 20-day holding period. This tests whether the "purity" of the 9-count setup leads to more sustained trends compared to the potentially premature 8-count.

Table 3: Comparative Performance Metrics for $k = 20$ (Medium-Term Trend Change)

	9-4 Pattern				8-3 Pattern				Statistical Tests	
Asset	**PF**	**Hit %**	**Exp.**	**Eff.**	**PF**	**Hit %**	**Exp.**	**Eff.**	**T-Test**	**KS-Test**
EURUSD	0.92	52.81%	-13.53	-0.68	0.91	51.48%	-15.05	-0.75	0.9683	0.8896
USDCHF	0.88	49.20%	-21.04	-1.05	1.00	52.92%	0.58	0.03	0.5649	0.5031
USDCAD	1.00	48.30%	0.14	0.01	0.90	48.47%	-10.51	-0.53	0.6958	0.9925
EURSEK	1.13	53.67%	14.52	0.73	1.10	52.10%	11.61	0.58	0.9227	0.9025
CHFJPY	0.78	44.82%	-37.00	-1.85	1.15	50.18%	21.41	1.07	0.0801	0.3602
EURHUF	1.10	52.31%	10.27	0.51	1.09	47.85%	10.54	0.53	0.9929	0.4536
USDMXN	1.07	52.50%	16.57	0.83	1.04	55.06%	7.79	0.39	0.8800	0.9891
USDSGD	0.90	48.82%	-10.19	-0.51	0.85	50.70%	-15.29	-0.76	0.8140	0.9837

Note: The only asset approaching statistical significance is CHFJPY ($p = 0.08$), where the Accelerated (8) setup actually outperformed the Canonical (9) setup (Profit Factor 1.15 vs 0.78). For all other pairs, the null hypothesis holds firmly.

4.5 Overall Interpretation

Across all three time horizons ($k = 5, 10, 20$), the data reveals a consistent pattern: the performance of the Alternative Setup (8, 3) is statistically indistinguishable from the Canonical Setup (9, 4). The high p-values across the T-Tests and KS-Tests indicate that the specific integer constraint of "9" does not possess unique predictive alpha compared to "8". In instances where divergence occurred (e.g., CHFJPY at $k = 20$), the (8, 3) setup occasionally exhibited superior metrics, further challenging the rigidity of the traditional parameters.