2. Empirical Analysis: 30-City Experiment

2.1. Descriptive statistics

Table 1 Descriptive Statistics by Region Status (Swing / Win / Lose)

Metric	Status	N	Mean	Std	Min	Median	Q3	Max
Daily Wage	Lose	2623	1369.09	914.11	616.50	1222.20	1984.40	5196.45
	Swing	1359	1319.05	760.51	694.80	1231.38	1893.86	4037.28
	Win	1150	1385.02	776.65	768.12	1307.70	1958.28	4675.47
Daily Hours Worked	Lose	2623	4.5544	2.9845	2.0522	4.1808	6.6450	20.7169
	Swing	1359	5.1528	2.9436	2.7818	4.9550	7.2388	16.9994
	Win	1150	5.5401	3.0070	3.1342	5.3775	7.7347	15.6217
Daily Orders	Lose	2623	20.73	13.89	9.0	18.0	30.0	77.0
	Swing	1359	25.06	14.65	13.0	23.0	36.0	71.0
	Win	1150	27.58	15.42	15.0	26.0	40.0	81.0
Days Worked	Lose	2623	3.376	0.859	3.0	4.0	4.0	4.0
	Swing	1359	3.327	0.862	3.0	4.0	4.0	4.0
	Win	1150	3.356	0.847	3.0	4.0	4.0	4.0

We classify the 30-city region into three performance-based segments: ”Win”, ”Swing” and ”Lose”. ”Win” regions denote regions that outperform all rivals. ”Swing” regions are those characterized by a narrow margin between the subject platform and its primary competitors. ”Lose” regions consist of territories where the subject platform trails behind the leading providers. Each row reports the total number of drivers (n), mean, standard deviation (Std), minimum (Min), median, third quartile (Q3), and maximum (Max) for workers in each region category. Differences in daily wages and order volumes suggest that workers in winning regions exhibit moderately higher productivity and earnings compared to swing and losing regions.

Table 2 Descriptive Statistics by Region Tier (Bottom / Mid / Top)

Metric	Region	n	Mean	Std	Min	Median	Q3	Max
Daily Wage	Bottom	1824	914.48	689.61	394.29	742.32	1267.42	4466.16
	Mid	1917	1311.64	765.54	721.80	1206.18	1811.88	5196.45
	Top	1390	1854.30	817.07	1273.68	1874.70	2399.88	5083.11
Daily Hours Worked	Bottom	1824	3.3311	2.4897	1.3928	2.7468	4.7449	20.7169
	Mid	1917	4.7382	2.6920	2.6386	4.4847	6.5544	16.9994
	Top	1390	6.7406	2.8566	4.7942	6.8308	8.7250	15.7308
Daily Orders	Bottom	1824	15.35	11.73	6.0	12.0	22.0	70.0
	Mid	1917	22.64	13.41	12.0	21.0	31.0	70.0
	Top	1390	32.24	14.03	22.0	32.0	43.0	81.0
Days Worked	Bottom	1824	3.064	0.976	2.0	3.0	4.0	4.0
	Mid	1917	3.353	0.831	3.0	4.0	4.0	4.0
	Top	1390	3.655	0.634	3.0	4.0	4.0	4.0

We classify drivers into three separate categories: ”Bottom”, ”Mid”, and ”Top”. ”Bottom” drivers are drivers that perform the least compared to other drivers. ”Mid” drivers are drivers that perform at and average level, and ”Top” drivers are top-performing drivers. Each row reports the total number of drivers (n), mean, standard deviation (Std), minimum (Min), median, third quartile (Q3),and maximum (Max). Top workers earn substantially higher daily wages (mean = 1854.30) and fulfill more daily orders (mean = 32.24) compared to Bottom workers (mean = 914.48 and 15.35, respectively), indicating the clear distinction between driver tiers and productivity.

Table 3: Descriptive Statistics by Region and Metric

Region	**Status**	**Metric**	**n**	**Mean**	**Std**	**Median**
Chanthaburi	Lose	Daily Wage	133	1624.02	883.41	1637.90
		Daily Hours Worked	133	6.19	3.43	6.12
		Daily Orders	133	29.57	16.15	30.0
		Days Worked	133	3.239	0.904	3.5
Chiangrai	Lose	Daily Wage	387	1088.24	806.11	890.40
		Daily Hours Worked	387	3.68	2.74	3.00
		Daily Orders	387	17.32	12.98	14.0
		Days Worked	387	3.288	0.900	4.0
Chumphon	Win	Daily Wage	117	1458.42	669.60	1531.26
		Daily Hours Worked	117	7.07	3.13	7.32
		Daily Orders	117	35.52	16.57	37.0
		Days Worked	117	3.407	0.823	4.0
Hat Yai	Lose	Daily Wage	516	1204.38	805.01	1075.32
		Daily Hours Worked	516	4.28	2.75	3.92
		Daily Orders	516	18.85	12.98	16.0
		Days Worked	516	3.398	0.849	4.0
Kanchanaburi	Swing	Daily Wage	156	1233.63	728.87	1128.90
		Daily Hours Worked	156	4.52	2.60	4.27
		Daily Orders	156	21.83	13.00	20.0
		Days Worked	156	3.366	0.838	4.0
Khonkaen	Lose	Daily Wage	596	1472.29	925.87	1369.08
		Daily Hours Worked	596	5.09	3.22	4.71
		Daily Orders	596	22.74	14.48	21.0
		Days Worked	596	3.331	0.882	4.0
Krabi	Win	Daily Wage	271	1330.16	724.74	1252.20

Continued on next page →

Table 3 — continued from previous page

Region	Status	Metric	n	Mean	Std	Median
		Daily Hours Worked	271	5.10	2.68	4.99
		Daily Orders	271	27.22	14.88	25.0
		Days Worked	271	3.410	0.815	4.0
Lampang	Swing	Daily Wage	202	1282.79	724.70	1242.30
		Daily Hours Worked	202	5.27	3.06	5.17
		Daily Orders	202	23.96	13.97	22.0
		Days Worked	202	3.303	0.881	4.0
Nakhonnayok	Win	Daily Wage	63	1463.80	762.00	1487.97
		Daily Hours Worked	63	5.52	2.88	5.47
		Daily Orders	63	27.48	14.21	27.0
		Days Worked	63	3.478	0.782	4.0
Nakhonpathom	Lose	Daily Wage	382	1661.85	1108.75	1448.28
		Daily Hours Worked	382	4.49	2.98	3.98
		Daily Orders	382	20.18	13.67	18.0
		Days Worked	382	3.550	0.719	4.0
Nakhonsawan	Lose	Daily Wage	93	1153.37	699.00	1111.50
		Daily Hours Worked	93	4.05	2.43	3.88
		Daily Orders	93	17.40	10.57	17.0
		Days Worked	93	3.372	0.906	4.0
Nakhon-srithammarat	Swing	Daily Wage	219	1394.23	786.82	1300.80
		Daily Hours Worked	219	5.61	3.11	5.30
		Daily Orders	219	26.92	15.74	25.0
		Days Worked	219	3.316	0.865	4.0
Nongkhai	Swing	Daily Wage	91	1401.98	797.63	1245.12
		Daily Hours Worked	91	4.77	2.59	4.45
		Daily Orders	91	24.40	13.89	22.0

Continued on next page →

Table 3 — continued from previous page

Region	Status	Metric	n	Mean	Std	Median
		Days Worked	91	3.271	0.866	4.0
Phayao	Swing	Daily Wage	59	1431.88	802.84	1432.80
		Daily Hours Worked	59	5.59	3.02	5.88
		Daily Orders	59	29.31	16.05	30.0
		Days Worked	59	3.088	0.857	3.0
Phetchaburi (Chaam)	Swing	Daily Wage	63	1804.38	766.51	1815.51
		Daily Hours Worked	63	5.37	2.29	5.46
		Daily Orders	63	27.54	11.68	28.0
		Days Worked	63	3.589	0.724	4.0
Prachinburi	Swing	Daily Wage	157	1215.29	726.46	1118.09
		Daily Hours Worked	157	4.44	2.62	4.19
		Daily Orders	157	24.07	14.76	22.0
		Days Worked	157	3.270	0.935	4.0
Ratchaburi (Banpong)	Swing	Daily Wage	121	1348.07	799.67	1249.05
		Daily Hours Worked	121	5.54	3.33	5.07
		Daily Orders	121	25.82	15.50	23.0
		Days Worked	121	3.383	0.832	4.0
Sakaeo (Aranyaprathet)	Win	Daily Wage	81	1621.20	927.51	1477.80
		Daily Hours Worked	81	5.12	2.94	4.83
		Daily Orders	81	25.78	14.86	24.0
		Days Worked	81	3.303	0.934	4.0
Sakonnakhon	Win	Daily Wage	197	1383.04	722.88	1356.75
		Daily Hours Worked	197	6.07	3.11	6.03
		Daily Orders	197	28.77	14.95	28.0
		Days Worked	197	3.286	0.841	4.0

Continued on next page →

Table 3 — continued from previous page

Region	Status	Metric	n	Mean	Std	Median
Saraburi (Nongkhae)	Win	Daily Wage	76	1637.39	984.53	1502.13
		Daily Hours Worked	76	5.05	2.85	4.86
		Daily Orders	76	26.28	16.63	23.0
		Days Worked	76	3.218	0.926	3.0
Saraburi (Phraphutthabat)	Win	Daily Wage	34	1681.64	850.26	1584.72
		Daily Hours Worked	34	6.22	3.06	6.17
		Daily Orders	34	28.71	14.58	27.0
		Days Worked	34	3.519	0.826	4.0
Sisaket	Win	Daily Wage	156	1192.33	641.81	1154.78
		Daily Hours Worked	156	5.52	2.92	5.38
		Daily Orders	156	25.62	13.73	25.0
		Days Worked	156	3.286	0.857	4.0
Songkhla (City)	Lose	Daily Wage	154	1425.66	814.29	1338.84
		Daily Hours Worked	154	4.86	2.75	4.70
		Daily Orders	154	22.56	13.05	21.0
		Days Worked	154	3.366	0.877	4.0
Suratthani (City)	Lose	Daily Wage	78	1352.69	741.41	1319.52
		Daily Hours Worked	78	4.86	2.75	4.59
		Daily Orders	78	24.63	13.77	24.0
		Days Worked	78	3.346	0.875	4.0
Surin	Swing	Daily Wage	180	1179.43	678.15	1095.00
		Daily Hours Worked	180	5.51	3.07	5.31
		Daily Orders	180	26.29	15.18	25.0
		Days Worked	180	3.356	0.876	4.0
Tak (Maesot)	Swing	Daily Wage	111	1265.85	742.15	1225.89

Continued on next page →

Table 3 — continued from previous page

Region	**Status**	**Metric**	**n**	**Mean**	**Std**	**Median**
		Daily Hours Worked	111	4.88	2.82	4.66
		Daily Orders	111	23.48	14.14	22.0
		Days Worked	111	3.300	0.807	3.0
Trat	Win	Daily Wage	74	1369.28	871.68	1156.92
		Daily Hours Worked	74	4.94	3.12	4.22
		Daily Orders	74	25.34	16.06	21.5
		Days Worked	74	3.491	0.842	4.0
Udonthani	Lose	Daily Wage	227	1399.01	898.70	1247.82
		Daily Hours Worked	227	4.56	2.89	4.20
		Daily Orders	227	20.42	13.44	18.0
		Days Worked	227	3.388	0.868	4.0
Uttaradit	Lose	Daily Wage	57	994.69	646.37	925.20
		Daily Hours Worked	57	3.67	2.29	3.51
		Daily Orders	57	20.54	13.31	19.0
		Days Worked	57	3.239	0.924	4.0
Yasothon	Win	Daily Wage	81	1198.63	810.86	1127.15
		Daily Hours Worked	81	4.59	2.87	4.33
		Daily Orders	81	22.35	15.36	20.0
		Days Worked	81	3.263	0.840	3.0

Note: Status categories: *Win* = outperforms all rivals; *Swing* = narrow margin with primary competitors; *Lose* = trails leading providers. Daily wage is in local currency. Daily hours worked and days worked reflect weekly activity patterns. Daily orders indicate fulfillment volume.

Each row reports region, region status (Win/Swing/Lose), Metric for the statistics, the total number of drivers in the region(n), mean for each metric, and standard deviation (Std) for each metric.

2.2. Empirical strategy

The empirical objective is to test whether the *sequence* in which incentive types are assigned—not merely the current incentive—explains variation in market-level outcomes. We estimate a two-way fixed effects model in which the dependent variable is a city-week aggregate outcome, city effects absorb time-invariant heterogeneity across markets, and week effects absorb common calendar shocks. The key variation is the interaction between the current incentive assignment Z_{ct} and the lagged assignment $Z_{c,t-1}$.

Let $\bar{Y}_{ct}$ denote the mean per-driver outcome in city c during treatment week t: either mean daily online hours, mean daily completed orders, or mean daily acceptance rate. The interaction model is

$$\bar{Y}_{ct} = \alpha_c + \gamma_t + \sum_{k\in\mathcal{K}} \beta_k \mathbf{1}[Z_{ct} = k] + \sum_{j\in\mathcal{K}} \sum_{k\in\mathcal{K}} \lambda_{jk} \left(\mathbf{1}[Z_{c,t-1} = j] \cdot \mathbf{1}[Z_{ct} = k]\right) + \varepsilon_{ct} \tag{1}$$

where $\mathcal{K} = \{\mathrm{DH}, \mathrm{DO}, \mathrm{SH}, \mathrm{SO}\}$ and Control is the omitted reference category for both current and lagged incentives. The β_k coefficients capture the average level effect of each incentive relative to Control weeks, and the λ_{jk} coefficients capture the *incremental* effect of the specific predecessor j on the effectiveness of the current incentive k. The full 5×5 matrix of λ_{jk} values—the transition matrix—is the primary object of inference. Diagonal cells (same incentive in consecutive weeks) are structurally empty because the design prohibits back-to-back repetition of the same incentive, leaving $4 \times 4 = 16$ estimable interaction cells, with 12 observed after excluding the diagonal.

Standard errors are clustered at the city level throughout. All regressions are estimated with `linearmodels.PanelOLS` using entity and time fixed effects.

Three-level hypothesis structure. We test sequencing effects at three successively refined levels, following a pre-specified hierarchy to control for multiple comparisons.

Level 1. A joint F-test of $H_0\colon \lambda_{jk} = 0 \ \forall\, j, k$ evaluates whether the transition matrix contains any non-zero entry. The test statistic uses $q = 12$ degrees of freedom (12 observed non-diagonal interactions). A significant Level 1 result means that knowing the specific sequence $(j \to k)$ predicts the outcome beyond knowing only the current incentive k.

Level 2. Conditional on a significant Level 1 result, we run four row-wise F-tests, one for each current incentive $k \in \mathcal{K}$, testing whether the predecessor makes a difference within that column of the transition matrix. Bonferroni correction is applied: the threshold per test is $\alpha^* = 0.05/4 = 0.0125$.

Level 3. For rows that pass Level 2, we conduct all pairwise comparisons $\lambda_{jk} = \lambda_{j'k}$ among predecessors. Benjamini–Hochberg (BH) correction is applied jointly across all contrasts from all passing rows. BH controls the false discovery rate and is less conservative than Bonferroni for discovery-oriented testing.

2.3. Main effects

Before examining sequencing interactions, Table 4 reports the main effects $\hat{\beta}_k$—the average impact of each incentive type relative to no-incentive weeks, pooling all 30 cities and all transition contexts.

Table 4 Do Incentives Boost Driver Activity? Estimated Effects vs. No-Incentive Weeks (All 30 Cities)

Incentive Type	Online Hours *change (p-value)*	Daily Orders *change (p-value)*	Acceptance Rate *change in pp (p-value)*
Daily Hours	+0.005 (0.161)	+0.012 (0.192)	+1.3 (0.536)
Daily Orders	+0.002 (0.430)	+0.012 (0.067)	**+3.2** (**0.003**)
Streak Hours	+0.002 (0.367)	+0.015 (0.098)	**+3.5** (**0.030**)
Streak Orders	+0.000 (0.947)	+0.031 (0.073)	**+6.2** (**0.008**)

Each row shows the average change in that outcome during an incentive week compared to a no-incentive week, after absorbing city and week fixed effects. Acceptance Rate is in percentage points (pp). **Bold** = statistically significant ($p < 0.05$). No incentive type reliably increased hours online or orders completed across all 30 cities. All three order-completion incentive types (Daily Orders, Streak Hours, Streak Orders) significantly raised acceptance rates by 3–6 pp.

The absence of significant hours and orders effects at the city-week level is consistent with the heterogeneity documented below: incentive effects on volume outcomes are concentrated within specific market tiers, and averaging across Win, Swing, and Lose markets attenuates them toward zero.

2.4. Sequencing interactions

Level 1: Global test. Table 5 reports the Level 1 joint F-tests by market group and outcome.

Table 5 Does the Order of Incentives Matter? Joint F-Test Summary by Market Type

Market Group	Online Hours	Daily Orders	Acceptance Rate
All 30 Cities	No ($p = 0.059$)	No ($p = 0.697$)	No ($p = 0.586$)
Win Cities	**YES** ($p < 0.001$)	No ($p = 0.736$)	No ($p = 0.347$)
Swing Cities	**YES** ($p < 0.001$)	**YES** ($p < 0.001$)	**YES** ($p < 0.001$)
Lose Cities	Borderline ($p = 0.068$)	**YES** ($p < 0.001$)	No ($p = 0.543$)

YES indicates that the joint F-test on all 12 non-diagonal λ_{jk} terms rejects $H_0\colon \lambda_{jk} = 0\ \forall j,k$ at the 5% level. Win/Swing/Lose reflect market competitive tiers (10 cities each). Sequencing effects are strongest in Swing cities, where all three outcomes are affected. No sequencing effect is detected in the pooled 30-city sample, consistent with effects that vary in sign across market tiers and cancel when aggregated. Transition matrices for all **YES** cells appear in Tables 6–10.

The full 30-city sample shows no significant interactions for any outcome after pooling Win, Swing, and Lose markets. Within tiers, however, significant sequencing effects emerge. Swing cities show interactions on all three outcomes; Win cities show effects on online hours; Lose cities show effects on daily orders. The pattern suggests that sequencing operates differently across market environments and that aggregation masks heterogeneous treatment effects.

Levels 2 and 3: Transition matrices. Tables 6–10 present the 5×4 transition matrices for each significant cell from Table 5. Rows are the predecessor incentive; columns are the current incentive. Each cell reports the total estimated effect $\hat{\beta}_k + \hat{\lambda}_{jk}$ relative to a no-incentive week. Bold (italic) cells indicate predecessors that are significantly better (worse) than at least one alternative predecessor in the same column, after BH correction.

Table 6 Win Cities: Total Effect on Daily Online Hours by Incentive Sequence (hrs/driver/day)

	This Week's Incentive			
Last Week's Incentive	Daily Hours	Daily Orders	Streak Hours	Streak Orders
No Incentive	+0.008	+0.006	−0.001	−0.006
Daily Hours	—	+0.006	−0.008	+0.003
Daily Orders	+0.008	—	*−0.013*	+0.003
Streak Hours	−0.004	**+0.004**	—	+0.000
Streak Orders	−0.004	*−0.010*	**+0.016**	—

Values show the expected change in average daily online hours per driver vs. a no-incentive week. **Bold** = significantly better predecessor for that column; *italic* = significantly worse predecessor (pairwise comparisons, BH-adjusted $p \leq 0.05$). "No Incentive" row is the reference; "—" = same consecutive incentive, structurally excluded. The Streak Orders → Streak Hours sequence yields the highest online-hours uplift (+0.016 hrs); Daily Orders → Streak Hours yields the lowest (−0.013 hrs), a 0.029-hr difference from the same current incentive. Note: absolute magnitudes are small (under 2 minutes per driver per day).

Table 7 Swing Cities: Total Effect on Daily Online Hours by Incentive Sequence (hrs/driver/day)

	This Week's Incentive			
Last Week's Incentive	Daily Hours	Daily Orders	Streak Hours	Streak Orders
No Incentive	+0.002	+0.001	+0.001	−0.003
Daily Hours	—	+0.005	+0.006	+0.005
Daily Orders	+0.011	—	+0.005	+0.008
Streak Hours	+0.011	+0.006	—	+0.011
Streak Orders	+0.017	+0.008	+0.006	—

No individual predecessor was confirmed significantly better or worse at the BH 5% threshold. Nevertheless, the overall sequencing test is significant ($p < 0.001$): cities with any active incentive the prior week show systematically higher online hours during Streak Orders weeks (+0.005 to +0.011 hrs) compared to cities entering from a no-incentive week (−0.003 hrs). The *type* of prior incentive is less important than simply having had one.

Table 8 Swing Cities: Total Effect on Daily Orders by Incentive Sequence (orders/driver/day)

	This Week's Incentive			
Last Week's Incentive	Daily Hours	Daily Orders	Streak Hours	Streak Orders
No Incentive	+0.012	+0.007	+0.018	+0.045
Daily Hours	—	+0.018	+0.013	**+0.074**
Daily Orders	+0.031	—	*+0.035*	−0.002
Streak Hours	+0.035	+0.064	—	*+0.038*
Streak Orders	+0.040	+0.072	**+0.113**	—

Bold = significantly better predecessor; *italic* = significantly worse (BH-adjusted $p \leq 0.05$). The Streak Orders → Streak Hours sequence delivers the largest order uplift (+0.113 orders/day), roughly 6× the no-predecessor baseline (+0.018). Daily Hours → Streak Orders is also strong (+0.074). Conversely, Daily Orders is the weakest predecessor for a Streak Hours week and nearly eliminates the Streak Orders effect altogether (−0.002 vs. +0.045 baseline).

Table 9 Swing Cities: Total Effect on Acceptance Rate by Incentive Sequence (percentage points)

	This Week's Incentive			
Last Week's Incentive	Daily Hours	Daily Orders	Streak Hours	Streak Orders
No Incentive	+4.5	+2.9	+5.5	+9.1
Daily Hours	—	+2.2	+6.1	**+22.5**
Daily Orders	+4.4	—	+5.3	*−0.7*
Streak Hours	*+0.6*	+14.0	—	*+14.0*
Streak Orders	**+6.6**	+12.8	+7.6	—

Bold = significantly better predecessor; *italic* = significantly worse (BH-adjusted $p \leq 0.05$). The Daily Hours → Streak Orders sequence generates a +22.5 pp increase in acceptance rate, more than twice the no-predecessor baseline of +9.1 pp. The same current incentive (Streak Orders) following Daily Orders reduces acceptance rates below the no-incentive baseline (−0.7 pp). The 23 pp swing from the best to the worst predecessor for Streak Orders demonstrates that the choice of prior incentive is as consequential as the current incentive itself.

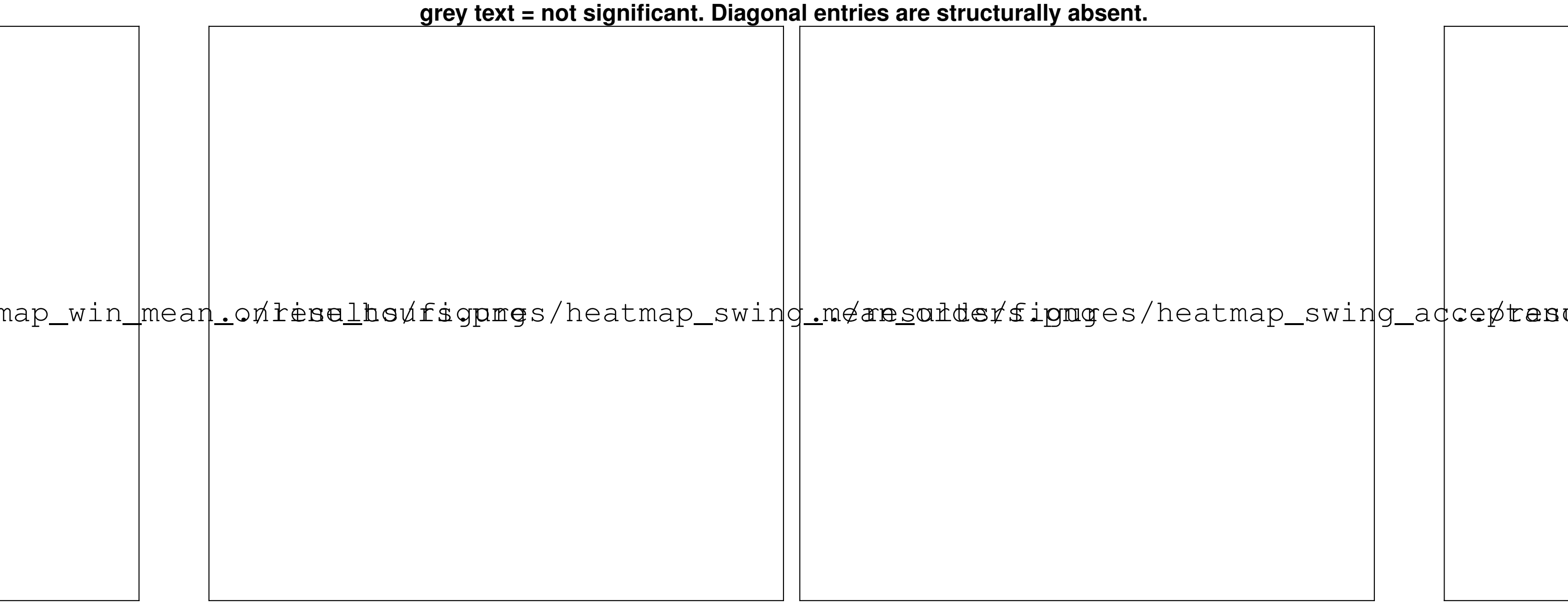
Table 10 Lose Cities: Total Effect on Daily Orders by Incentive Sequence (orders/driver/day)

	This Week's Incentive			
Last Week's Incentive	Daily Hours	Daily Orders	Streak Hours	Streak Orders
No Incentive	+0.006	+0.014	+0.026	+0.034
Daily Hours	—	+0.001	+0.057	+0.053
Daily Orders	*−0.045*	—	+0.017	+0.015
Streak Hours	**+0.045**	*+0.031*	—	+0.044
Streak Orders	*+0.025*	**+0.065**	+0.033	—

Bold = significantly better predecessor; *italic* = significantly worse (BH-adjusted $p \leq 0.05$). In Lose cities, the wrong predecessor can actively reduce outcomes below baseline. Daily Orders → Daily Hours yields a *negative* effect (−0.045 orders/day), worse than no incentive at all. Streak Hours → Daily Hours delivers +0.045 orders/day: a 0.090-order swing from the same current incentive. Streak Orders is the best predecessor for Daily Orders weeks (+0.065 vs. +0.031 for the second-best).

Heatmap visualization. Figure 2 presents the full λ_{jk} coefficient matrices as heatmaps for each market tier and outcome. Blue cells indicate positive interactions (the predecessor amplifies the current incentive above its baseline level effect); red cells indicate negative interactions. Black text marks cells that are statistically significant after BH correction at $\alpha = 0.05$; grey text marks insignificant cells. Diagonal entries are structural zeros.

Figure 2 Transition matrix heatmaps for statistically significant market-outcome pairs. Rows = predecessor incentive; columns = current incentive. Blue = positive λ_{jk}; red = negative. Black text = BH-significant at 5%; grey text = not significant. Diagonal entries are structurally absent.



(b) Swing: daily orders

(c) Swing: acceptance rate

Note. Each panel corresponds to a market-tier/outcome combination where the Level 1 joint F-test rejected the null of no sequencing effects. The Swing acceptance-rate panel (bottom left) exhibits the widest range of λ_{jk} values and the most significant predecessor contrasts, including a 23-percentage-point swing between the best and worst predecessor for the Streak Orders column.

2.5. Validation

Permutation test. To verify that the F-test p-values are reliable with only 30 clusters, we implement a randomisation-based test. In each of 1,000 replications, incentive labels are shuffled within the same competitive-tier × week cell, preserving the design constraint that exactly two cities per tier are assigned to each incentive in each treatment week. The partial-F statistic is recomputed under each shuffle, generating a null distribution free of distributional assumptions.

Table 11 **Randomisation Test and Stability Check**

Panel A: Randomisation Test

Outcome	Observed F	95th pct (random)	Conclusion
Online Hours	0.97	1.89	Consistent with chance
Daily Orders	1.31	4.47	Consistent with chance
Acceptance Rate	0.91	3.89	Consistent with chance

Panel B: Stability Test

	Online Hours	Daily Orders	Acceptance Rate
Maximum change ratio (threshold: <1.0)	0.34	0.40	0.80
All 12 sequences pass?	**Yes**	**Yes**	**Yes**
Sequencing effects are genuine — not explained by prior-week momentum.			

Panel A: The full-sample test statistic falls within the range of outcomes under random assignment for all three outcomes, confirming that the absence of overall effects across 30 cities is not a power problem but reflects genuine cancellation across market tiers. *Panel B*: Change ratio = $|\hat{\lambda}_{\text{base}} - \hat{\lambda}_{\text{controlled}}|/\text{SE}(\hat{\lambda}_{\text{base}})$, where the controlled model adds last week's realized outcome as a covariate. Ratios below 1.0 for all 12 sequences confirm that sequencing effects do not merely reflect mean reversion from unusually high or low prior-week performance.

Panel A confirms that the full-sample null findings are genuine: the observed statistics lie well within the null distribution, ruling out size distortion from the small number of clusters. The significant within-tier results (Table 5) arise because sequencing effects are real in each tier but differ in sign, cancelling out in the aggregate. Panel B shows that adding the lagged outcome to the model changes none of the interaction estimates by more than one standard error, ruling out mean reversion as an alternative explanation.

3. Empirical Analysis: Bangkok

4. Switchback Experiment

5. Setting: On-Demand Food Delivery Platform in Thailand

We study supply management on a leading on-demand food delivery platform in Thailand. The platform operates in a setting where demand is sharply time-varying, service quality depends on acquiring rider capacity at the right times and in the right places, and workers retain substantial discretion over whether to work, how long to remain active, and which assignments to accept. Incentive contracts are therefore one of the platform's central operational instruments. The question is not merely whether incentives increase labor supply, but how contracts should be structured when worker response is multi-margin and when current incentives can shape the response to future ones.

Thailand is a useful setting for studying this problem because the underlying operating environment closely resembles major U.S. delivery markets. The sector is organized around multi-sided digital platforms that match consumers and restaurants with a flexible fleet of independent riders. Demand exhibits strong meal-time peaks and is further perturbed by weather, traffic, and local events. Riders self-schedule and make repeated short-horizon decisions throughout the day, including whether to log on, how long to remain active, and whether to accept particular assignments. As in other delivery markets, the platform must manage capacity under pronounced temporal and spatial demand variation.

At the same time, the Thai setting sharpens the interpretation of incentive effects. Deliveries are predominantly fulfilled via motorbikes in dense urban environments, which heightens the operational importance of route familiarity, travel-time uncertainty, and weather exposure. The setting also appears to feature less seamless within-day multihoming than many U.S. gig markets. Riders commonly operate with platform-branded uniforms and insulated delivery boxes whose design and mounting can be platform-specific, and platform-specific routines around onboarding, conduct, and operations can further increase the cost of switching mid-shift. These frictions strengthen the mapping from platform policy to realized within-platform supply. In such an environment, incentive contracts are less likely to simply reallocate attention across simultaneously open apps and more likely to shape actual participation, persistence, and acceptance behavior on the focal platform (Thairath Money 2024, Momentum Works 2024).

5.1. Platform operations and rider decisions

The platform operates a standard dispatch-based workflow. When a rider is active, the platform surfaces delivery opportunities and assigns tasks through a combination of algorithmic matching and operational rules. Riders receive a stream of assignments, each associated with pickup and drop-off locations, compensation, and expected effort, and then decide whether to accept. Upon acceptance, the rider executes pickup and delivery, after which the platform records completion time, distance, and compensation components. This workflow makes rider behavior inherently multi-margin. A rider first decides whether to work at all. Conditional on being active, the rider chooses how long to remain online. Conditional on remaining online, the rider also chooses how selective to be in accepted work. These margins map directly into platform performance. A larger set of active riders expands the labor pool. Longer online duration increases service coverage. Lower selectivity can raise realized throughput, but may also affect travel distance, waiting time, and service quality. The platform's incentive problem is therefore not simply to increase labor input in the aggregate, but to shape the mix of participation, availability, and acceptance behavior that best supports service performance.

5.2. Contract primitives in the field setting

The platform's operational incentive menus can be organized around two dimensions. The first is the *qualification metric*. Some contracts reward *availability*, typically by requiring riders to remain online for a sufficient number of hours. Others reward *throughput*, typically by requiring riders to complete a threshold number of orders. The second is the *reward horizon*. Some contracts *reset* at the end of the day. Others are *state-dependent*, so that qualifying today changes the value of future participation within the week.

These dimensions are operationally meaningful. An hours-based contract is intended to purchase presence. An orders-based contract is intended to purchase conversion conditional on participation. A reset contract targets current behavior, whereas a state-dependent contract creates intertemporal incentives to maintain participation and preserve progress toward later payout.

In practice, these contract families are implemented through several concrete menus. Availability-based contracts often combine a total-hours requirement with a peak-hours requirement, so that availability is supplied when it is operationally most valuable rather than uniformly over the day. Throughput-based contracts reward completed orders rather than mere presence and therefore interact more directly with rider selectivity and local market conditions. State-dependent contracts take two main forms in our field setting. One is a weekly hours scheme in which riders must qualify on a sufficient number of days and, in some cases, on mandatory high-demand days. The other is a segmented weekly output scheme in which the week is partitioned into several windows and riders earn bonuses for satisfying output requirements throughout each segment. In both cases, current qualification changes the value of future participation because it changes the rider's progress toward later payout.

The platform also conditions thresholds and bonus amounts on rider performance tiers. Riders are grouped into broad categories based on recent activity and performance, and the platform posts tier-specific targets so that qualification remains attainable for less active riders while remaining operationally meaningful for highly active riders. These institutional features yield a natural contract space organized around the metric used to determine qualification and the horizon over which qualification is rewarded. Those are the primitives formalized in the model and, in turn, the dimensions that organize the experimental variation.

6. Experimental Design

The theoretical framework implies that an informative experiment must vary the two contract primitives that govern worker response, namely the qualification metric and the reward horizon, and must do so in a way that permits predecessor-sensitive comparisons. The design must also respect the marketplace nature of the environment. Riders within the same local market compete for a common pool of orders, so rider-level randomization would alter matching conditions and contaminate treatment effects. For these reasons, the experiment was implemented as a market-level switchback in which incentive menus rotated across regions and weeks.

The platform implemented the experiment in 30 expansion regions outside Bangkok. The unit of assignment is a region-week. In each region-week (c,t), the platform posts a single incentive menu that applies to all eligible riders operating in region c during week t. This assignment level matches the platform's actual decision problem, namely which contract should be posted in a given market and week, while also mitigating within-market interference that would arise under individual-level randomization.

The experiment ran over two consecutive months, July 2025 and August 2025. In each month, the schedule followed a repeated three-week cadence. Week 1 was a recovery week with no region-specific weekly incentive. Weeks 2 and 3 were incentive weeks, during which each region was assigned to one of the four treatment families or to a no-incentive baseline. The two incentive weeks within each month therefore generate the core sequencing variation through the ordered pair of assignments in Weeks 2 and 3.

Before implementation, the platform grouped the 30 regions into three broad categories based on historical performance and competitive conditions: ten *win* markets, ten *swing* markets, and ten *lose* markets. This segmentation was operationally meaningful for the platform and also useful experimentally because it allowed assignment to be balanced within broad market-condition strata. In every incentive week, within each of the three groups, exactly two regions were assigned to each of the five conditions. Aggregating across groups, each condition appeared in six regions per incentive week. Formally, if C_g denotes the set of regions in cluster $g \in \{\text{win}, \text{swing}, \text{lose}\}$, then for every incentive week t and every condition $z \in \{\text{DH}, \text{DO}, \text{SH}, \text{SO}, 0\}$,

$$\sum_{c \in C_g} \mathbb{I}\{Z_{ct} = z\} = 2, \qquad \sum_{c=1}^{30} \mathbb{I}\{Z_{ct} = z\} = 6.$$

This structure serves two purposes. First, it sharply reduces dependence on functional-form assumptions about common calendar shocks, because identification comes from contrasts within the same week as well as within-region changes over time. Second, it creates replication of predecessor-successor transitions across distinct market environments, which improves precision and permits heterogeneity analyses by market strength.

The treatment menu follows directly from the contract space in the model. We vary two dimensions: whether qualification rewards *availability* or *throughput*, and whether rewards *reset* immediately or are *state-dependent* within the week. Crossing these dimensions yields four treatment families:

$$\{\text{DH}, \text{DO}, \text{SH}, \text{SO}\},$$

where DH denotes Daily Hours, DO denotes Daily Orders, SH denotes Streak Hours, and SO denotes Streak Orders. In addition, some region-weeks were assigned to a no-incentive baseline, denoted by 0, meaning no region-specific weekly incentive beyond business-as-usual platform adjustments. The object of inference is therefore the effect of contract *structure*, not merely the effect of higher payout.

The four treatment families were implemented through the platform's standard incentive-posting system. In practice, thresholds and payouts varied by rider performance tier. Riders were classified into three tiers, which we refer to as low, mid, and top, based on recent activity and performance. The platform used tier-specific targets so that qualification remained attainable for lower-activity riders while remaining operationally meaningful for highly active riders.

Let H_{icdt} denote rider i's total online hours on day d of week t in city c, and let H_{icdt}^{peak} denote online hours that fall within platform-defined peak windows. Let Q_{icdt} denote completed orders on that day. Under Daily Hours (DH), riders qualified for a daily bonus by satisfying both a total-hours requirement and a peak-hours requirement:

$$1\{\text{Qualify}_{icdt}^{DH} = 1\} = 1\{H_{icdt} \geq \bar{H}_{g(i,c,t)}\} \cdot 1\{H_{icdt}^{\text{peak}} \geq \bar{H}_{g(i,c,t)}^{\text{peak}}\}.$$

Under Daily Orders (DO), riders qualified for a daily bonus by completing at least a tier-specific order threshold:

$$1\{\text{Qualify}_{icdt}^{DO} = 1\} = 1\{Q_{icdt} \geq \bar{Q}_{g(i,c,t)}\}.$$

The two state-dependent treatments preserve the same distinction between hours-based and orders-based qualification while introducing within-week continuation value. Under Streak Hours (SH), riders had to satisfy the hours requirement on at least a specified number of days within the week and, in some implementations, also satisfy the requirement on mandatory high-demand days. Let $\mathcal{M}$ denote the mandatory-day set and let K_g denote the required number of qualifying days for tier g. Weekly qualification is then

$$1\{\text{Earn}_{ict}^{SH} = 1\} = 1\left\{\sum_{d=1}^{D} \text{Qualify}_{icdt}^{DH} \geq K_{g(i,c,t)}\right\} \cdot \prod_{d \in \mathcal{M}} \text{Qualify}_{icdt}^{DH}.$$

Under Streak Orders (SO), the week was partitioned into several segments, and riders earned bonuses for satisfying the daily output requirement on every day within a given segment. Let $\{\mathcal{S}_r\}_{r=1}^{R}$ denote the segment day sets. Then segment-r qualification is

$$1\{\text{Earn}_{ict}^{SO,r} = 1\} = \prod_{d \in \mathcal{S}_r} \text{Qualify}_{icdt}^{DO},$$

and total weekly bonus is the sum across completed segments. In the baseline condition, riders received the platform's standard compensation without additional region-specific weekly incentives.

These implementation details matter because they clarify how the abstract contract families map into the field intervention. In particular, SH and SO are not simply longer-horizon versions of DH and DO. They are concrete state-dependent contracts with mandatory-day and segmented-progress features that generate continuation value within the week.

Table 12 reports the full region-by-week assignment schedule. Two features are worth emphasizing. First, every incentive week contains all five conditions. Second, the two incentive weeks within each month generate predecessor-successor transitions that are replicated across win, swing, and lose markets.

Table 12 Region-by-week incentive assignment schedule in the 30-city experiment.

Region	Cluster	July 2025 W2 scheme	July 2025 W3 scheme	August 2025 W2 scheme	August 2025 W3 scheme
Saraburi (Phraphutthabat)	Win	○ Daily Hours	■ Streak Orders	• Streak Hours	— No Incentive
Chumphon	Win	○ Daily Hours	• Streak Hours	○ Daily Hours	— No Incentive
Nakhonnayok	Win	■ Streak Orders	□ Daily Orders	□ Daily Orders	• Streak Hours
Sakonnakhon	Win	■ Streak Orders	— No Incentive	■ Streak Orders	• Streak Hours
Saraburi (Nongkhae)	Win	□ Daily Orders	○ Daily Hours	— No Incentive	□ Daily Orders
Sisaket	Win	□ Daily Orders	— No Incentive	○ Daily Hours	□ Daily Orders
Trat	Win	• Streak Hours	■ Streak Orders	— No Incentive	■ Streak Orders
Sakaeo (Aranyaprathet)	Win	• Streak Hours	□ Daily Orders	□ Daily Orders	■ Streak Orders
Krabi	Win	— No Incentive	○ Daily Hours	• Streak Hours	○ Daily Hours
Yasothon	Win	— No Incentive	• Streak Hours	■ Streak Orders	○ Daily Hours
Nakhonsrithammarat	Swing	○ Daily Hours	■ Streak Orders	• Streak Hours	— No Incentive
Nongkhai	Swing	○ Daily Hours	• Streak Hours	○ Daily Hours	— No Incentive
Surin	Swing	■ Streak Orders	□ Daily Orders	□ Daily Orders	• Streak Hours
Tak (Maesot)	Swing	■ Streak Orders	— No Incentive	■ Streak Orders	• Streak Hours
Phayao	Swing	□ Daily Orders	○ Daily Hours	— No Incentive	□ Daily Orders
Lampang	Swing	□ Daily Orders	— No Incentive	○ Daily Hours	□ Daily Orders
Phetchaburi (Chaam)	Swing	• Streak Hours	■ Streak Orders	— No Incentive	■ Streak Orders
Prachinburi	Swing	• Streak Hours	□ Daily Orders	□ Daily Orders	■ Streak Orders
Kanchanaburi	Swing	— No Incentive	○ Daily Hours	• Streak Hours	○ Daily Hours
Ratchaburi (Banpong)	Swing	— No Incentive	• Streak Hours	■ Streak Orders	○ Daily Hours
Chanthaburi	Lose	○ Daily Hours	■ Streak Orders	• Streak Hours	— No Incentive
Suratthani (City)	Lose	○ Daily Hours	• Streak Hours	○ Daily Hours	— No Incentive
Nakhonsawan	Lose	■ Streak Orders	□ Daily Orders	□ Daily Orders	• Streak Hours
Khonkaen	Lose	■ Streak Orders	— No Incentive	■ Streak Orders	• Streak Hours
Udonthani	Lose	□ Daily Orders	○ Daily Hours	— No Incentive	□ Daily Orders
Songkhla (City)	Lose	□ Daily Orders	— No Incentive	○ Daily Hours	□ Daily Orders
Hat Yai	Lose	• Streak Hours	■ Streak Orders	— No Incentive	■ Streak Orders
Uttaradit	Lose	• Streak Hours	□ Daily Orders	□ Daily Orders	■ Streak Orders
Chiangrai	Lose	— No Incentive	○ Daily Hours	• Streak Hours	○ Daily Hours
Nakhonpathom	Lose	— No Incentive	• Streak Hours	■ Streak Orders	○ Daily Hours

Legend:

○ Daily Hours	(single-day, hours-based)	• Streak Hours	(multi-day, hours-based)
□ Daily Orders	(single-day, orders-based)	■ Streak Orders	(multi-day, orders-based)
— No Incentive			(control period)

The schedule was pre-specified and implemented through the platform's standard incentive-posting mechanism. We verify deployment using administrative records on posted menus, eligibility flags, and realized payouts, and we document deviations driven by operational exceptions. Two threats to interpretation are particularly relevant. The first is cross-region mobility. If riders frequently work across region boundaries, treatment effects may spill over across nominal assignment units. We therefore measure cross-region work patterns and later report robustness analyses that restrict attention to riders whose activity is concentrated in a single region during the relevant window. The second is dynamic anticipation and carryover. Because the design is intended to recover dynamic responses, we do not treat such persistence as a violation. Instead, the recovery weeks and predecessor-successor structure are used to estimate how current outcomes depend on recent contract history. The empirical strategy in §**??** combines fixed-effects estimation with randomization-based checks that respect the known assignment schedule.

7. Novelty Effect

A central challenge in gig-platform management is the ”decay” of driver engagement over time. We propose that driver response to incentives follows a distinct three-phase behavioral lifecycle: Arousal, Decay, and Recovery. This paper argues that the effectiveness of an incentive is not merely a function of its monetary value, but of its temporal architecture. By analyzing the ”Shock Factor” of new incentives and the non-linear decay of productivity throughout the week, we provide a framework for strategic incentive sequencing tailored to specific driver cohorts.

7.1. Statistical Modeling

To analyze the temporal dynamics of driver behavior, we employed a series of Linear Mixed-Effects Models (MLM). This approach was chosen to account for the nested structure of our data, as multiple daily observations are attributed to individual drivers, necessitating the inclusion of driver-specific random intercepts to control for unobserved individual heterogeneity ($Group\ Var$).

We first isolated the immediate behavioral response to incentive introduction by generating a binary indicator for the first day of each incentive period. Using No Incentive as the reference baseline, we estimated an interaction model:

$$Orders_{it} = \beta_0 + \beta_1(Day1_{it}) + \beta_2(Incentive_{it}) + \beta_3(Day1_{it} \times Incentive_{it}) + u_i + \epsilon_{it}$$

This model allowed us to determine if the ”Shock Factor” (the surge in productivity on Day 1)varied significantly across different incentive architectures.

We then used longitudinal decay and quadratic modeling to capture the evolution of productivity over the seven-day incentive cycle, performing separate longitudinal analyses for each incentive type. We introduced a quadratic term (Day^2) to the regression to account for non-nonlinearities in labor supply. A negative coefficient for the linear day term indicates a ”downward slide” or fatigue effect, while a positive quadratic coefficient identifies a late-week recovery or ”final sprint.”

Finally, we extended the model to include interaction terms between Rider Groups (Tiers) and the temporal variables (Day and Day^2). By setting the bottom tier as the baseline, we were able to observe how high-intensity drivers (Top Tier) differ from lower-intensity drivers in their susceptibility to initial shocks and their strategic use of late-week recovery phases.

Table 13 Mixed Linear Model Regression Results: Assessing the Day 1 Shock Factor

Variable	Coef.	Std.Err.	z	P> |z|	[0.025	0.975]
Intercept (Baseline)	19.404	0.134	145.28	0.000	19.142	19.666
Main Effects (Relative to No Incentive)						
Daily Hours	-3.798	0.088	-43.15	0.000	-3.970	-3.625
Daily Orders	-1.015	0.108	-9.35	0.000	-1.227	-0.802
Streak Hours	-0.932	0.084	-11.06	0.000	-1.098	-0.767
Streak Orders	-1.175	0.085	-13.84	0.000	-1.342	-1.009
Day 1 Shock Effects						
Is Day 1 (Baseline)	0.542	0.119	4.55	0.000	0.308	0.775
Is Day 1 × Daily Hours	1.863	0.203	9.16	0.000	1.464	2.261
Is Day 1 × Daily Orders	-0.176	0.230	-0.77	0.443	-0.626	0.274
Is Day 1 × Streak Hours	0.520	0.199	2.61	0.009	0.129	0.911
Is Day 1 × Streak Orders	-0.064	0.191	-0.33	0.739	-0.439	0.311
Group Variance (driver_id)	107.368	0.200				

Table 14 Decay Analysis for No Incentive

Variable	Coef.	Std.Err.	z	P> $\|z\|$	[0.025	0.975]
Intercept	21.887	0.222	98.788	0.000	21.453	22.321
Day of Incentive (Linear)	-0.870	0.103	-8.483	0.000	-1.071	-0.669
Day Squared (Quadratic)	0.093	0.013	7.018	0.000	0.067	0.119
Group Variance ($driver\_id$)	120.864	0.291				

Table 15 Decay Analysis for Streak Orders

Variable	Coef.	Std.Err.	z	P> $\|z\|$	[0.025	0.975]
Intercept	21.101	0.259	81.555	0.000	20.594	21.608
Day of Incentive	-1.731	0.105	-16.412	0.000	-1.938	-1.524
Day Sq.	0.231	0.013	17.511	0.000	0.205	0.257
Group Var	127.315	0.423				

Table 16 Decay Analysis for Daily Orders

Variable	Coef.	Std.Err.	z	P> $\lvert z \rvert$	[0.025	0.975]
Intercept	21.076	0.353	59.696	0.000	20.384	21.768
Day of Incentive	-1.127	0.150	-7.521	0.000	-1.421	-0.834
Day Sq.	0.133	0.019	7.170	0.000	0.097	0.170
Group Var	112.356	0.479				

Table 17 Decay Analysis for Streak Orders

Variable	Coef.	Std.Err.	z	P> $\lvert z \rvert$	[0.025	0.975]
Intercept	21.639	0.262	82.456	0.000	21.124	22.153
Day of Incentive	-1.065	0.118	-9.041	0.000	-1.296	-0.834
Day Sq.	0.099	0.015	6.750	0.000	0.070	0.127
Group Var	129.656	0.385				

Table 18 Decay Analysis for Daily Hours

Variable	Coef.	Std.Err.	z	P> $\lvert z \rvert$	[0.025	0.975]
Intercept	21.255	0.313	67.976	0.000	20.642	21.867
Day of Incentive	-1.192	0.145	-8.198	0.000	-1.476	-0.907
Day Sq.	0.059	0.018	3.269	0.001	0.023	0.094
Group Var	116.539	0.328				

7.2. Analysis

The initial phase of the incentive lifecycle is defined by a statistically significant surge in productivity, hereafter referred to as the ”Shock Factor.” Our analysis indicates that while a natural Day 1 baseline increase of 0.54 orders exists even in the absence of incentives, the introduction of targeted rewards creates a substantial ”jolt” to the system. Specifically, Daily Hours incentives seems to yield the most robust response, contributing an additional 1.86 orders above the control baseline for a total Day 1 surge of approximately 2.40 orders (this was smoothed out in the graph since we take the mean across all drivers with Daily Hours, that may have started thier incentive week on different calendar days). However, this initial boost was not sustained, sustaining a steep decay of -1.19 orders per day on average, the sharpest of all incentive types, resulting in a net loss of roughly 8

Table 19 Tiered Analysis for No Incentive

Variable	**Coef.**	**Std.Err.**	**z**	**P>** $\|z\|$	**[0.025**	**0.975]**
Intercept	17.140	0.352	48.726	0.000	16.450	17.829
C(Q('Rider Group'))[T.Mid]	5.778	0.456	12.672	0.000	4.884	6.672
C(Q('Rider Group'))[T.Top]	10.662	0.503	21.185	0.000	9.676	11.649
Day of Incentive	-0.886	0.190	-4.670	0.000	-1.258	-0.514
C(Q('Rider Group'))[T.Mid]:Day of Incentive	-0.105	0.252	-0.416	0.677	-0.599	0.389
C(Q('Rider Group'))[T.Top]:Day of Incentive	0.173	0.265	0.651	0.515	-0.347	0.692
Day Sq.	0.100	0.025	4.059	0.000	0.052	0.148
C(Q('Rider Group'))[T.Mid]:Day Sq.	0.006	0.033	0.173	0.863	-0.058	0.070
C(Q('Rider Group'))[T.Top]:Day Sq.	-0.035	0.034	-1.021	0.307	-0.102	0.032
Group Var	80.763	0.215				

Table 20 Tiered Analysis for Streak Orders

Variable	**Coef.**	**Std.Err.**	**z**	**P>** $\|z\|$	**[0.025**	**0.975]**
Intercept	15.751	0.379	41.536	0.000	15.008	16.494
C(Q('Rider Group'))[T.Mid]	6.627	0.484	13.685	0.000	5.678	7.576
C(Q('Rider Group'))[T.Top]	12.099	0.550	21.999	0.000	11.021	13.177
Day of Incentive	-0.992	0.190	-5.209	0.000	-1.365	-0.618
C(Q('Rider Group'))[T.Mid]:Day of Incentive	-0.885	0.255	-3.466	0.001	-1.386	-0.385
C(Q('Rider Group'))[T.Top]:Day of Incentive	-1.272	0.272	-4.686	0.000	-1.804	-0.740
Day Sq.	0.139	0.024	5.828	0.000	0.092	0.186
C(Q('Rider Group'))[T.Mid]:Day Sq.	0.104	0.032	3.262	0.001	0.042	0.167
C(Q('Rider Group'))[T.Top]:Day Sq.	0.167	0.034	4.905	0.000	0.100	0.233
Group Var	84.706	0.311				

orders relative to Day 1 productivity by the end of the week. This suggests that time-based targets serve as a more immediate behavioral trigger than volume-based targets, likely due to the higher salience and perceived controllability of labor hours at the start of a weekly cycle. However, this high-arousal state appears to be temporary, as the initial surge is followed by a period of declining marginal effort.

Table 21 Tiered Analysis for Daily Orders

Variable	**Coef.**	**Std.Err.**	**z**	**P> $\|z\|$**	**[0.025**	**0.975]**
Intercept	17.197	0.542	31.714	0.000	16.134	18.260
C(Q('Rider Group'))[T.Mid]	4.665	0.683	6.834	0.000	3.327	6.003
C(Q('Rider Group'))[T.Top]	9.130	0.751	12.163	0.000	7.659	10.601
Day of Incentive	-1.073	0.282	-3.805	0.000	-1.626	-0.520
C(Q('Rider Group'))[T.Mid]:Day of Incentive	-0.031	0.370	-0.085	0.932	-0.757	0.694
C(Q('Rider Group'))[T.Top]:Day of Incentive	-0.136	0.387	-0.350	0.726	-0.895	0.624
Day Sq.	0.133	0.035	3.769	0.000	0.064	0.201
C(Q('Rider Group'))[T.Mid]:Day Sq.	-0.007	0.046	-0.143	0.886	-0.097	0.084
C(Q('Rider Group'))[T.Top]:Day Sq.	0.015	0.048	0.315	0.753	-0.079	0.109
Group Var	77.379	0.363				

Table 22 Tiered Analysis for Streak Hours

Variable	**Coef.**	**Std.Err.**	**z**	**P> $\|z\|$**	**[0.025**	**0.975]**
Intercept	13.392	0.440	30.422	0.000	12.529	14.255
C(Q('Rider Group'))[T.Mid]	8.468	0.588	14.391	0.000	7.315	9.621
C(Q('Rider Group'))[T.Top]	19.240	0.623	30.860	0.000	18.018	20.462
Day of Incentive	-0.738	0.223	-3.315	0.001	-1.174	-0.301
C(Q('Rider Group'))[T.Mid]:Day of Incentive	-0.286	0.291	-0.983	0.326	-0.855	0.284
C(Q('Rider Group'))[T.Top]:Day of Incentive	-0.679	0.303	-2.244	0.025	-1.273	-0.086
Day Sq.	0.071	0.028	2.553	0.011	0.016	0.125
C(Q('Rider Group'))[T.Mid]:Day Sq.	0.027	0.036	0.737	0.461	-0.044	0.097
C(Q('Rider Group'))[T.Top]:Day Sq.	0.056	0.038	1.493	0.135	-0.018	0.130
Group Var	83.326	0.257				

Following the Day 1 peak, all incentive structures exhibit a universal Temporal Decay, as evidenced by the negative linear coefficients for the Day of Incentive variable in our MixedLM models. This downward trajectory represents the physiological and psychological cost of the initial effort, leading to a mid-week productivity trough. The divergence in how drivers finish the week, however, is determined by the architectural memory of the incentive. Streak-based incentives demonstrate a

Table 23 Tiered Analysis for Daily Hours

Variable	Coef.	Std.Err.	z	P> $\|z\|$	[0.025	0.975]
Intercept	17.269	0.528	32.686	0.000	16.234	18.305
C(Q('Rider Group'))[T.Mid]	4.535	0.681	6.659	0.000	3.200	5.870
C(Q('Rider Group'))[T.Top]	8.345	0.741	11.261	0.000	6.893	9.798
Day of Incentive	-0.762	0.270	-2.828	0.005	-1.291	-0.234
C(Q('Rider Group'))[T.Mid]:Day of Incentive	-0.763	0.354	-2.154	0.031	-1.457	-0.069
C(Q('Rider Group'))[T.Top]:Day of Incentive	-0.398	0.375	-1.060	0.289	-1.133	0.338
Day Sq.	0.068	0.033	2.036	0.042	0.003	0.133
C(Q('Rider Group'))[T.Mid]:Day Sq.	0.028	0.044	0.642	0.521	-0.058	0.114
C(Q('Rider Group'))[T.Top]:Day Sq.	-0.064	0.046	-1.384	0.166	-0.155	0.027
Group Var	105.189	0.304				

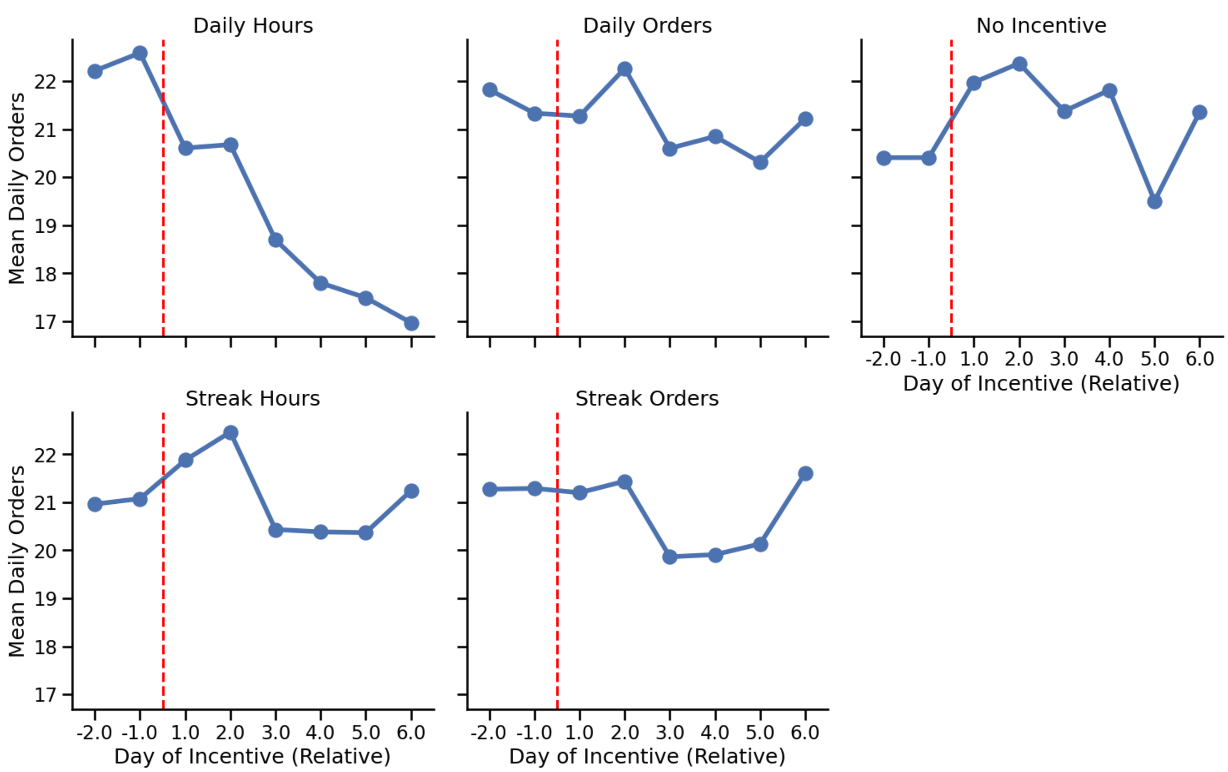


Figure 3 Mean Number of Orders Completed Throughout the Week Across All Drivers

strong "Goal Gradient Effect," where the cumulative nature of the reward acts as a persistence magnet. This is statistically captured by a positive quadratic term (day^2), which manifests as a "Final Sprint" or U-shaped productivity curve as the weekend approaches. The weekend-contingent nature of streak incentives creates a back-loaded motivation structure, effectively incentivizing drivers to maximize their labor supply at the end of the week to ensure the fulfillment of cumulative bonus requirements

In contrast, Daily level incentives function as ”memoryless” structures that lack a cumulative reward mechanism. Because a driver’s effort on Day 1 does not contribute to their eligibility for a Day 5 bonus, the mid-week slump often transitions into a state of permanent resignation rather than recovery. This effect is particularly pronounced among Top-Tier drivers, whose productivity curves fail to show the characteristic weekend upswing found in streak models. For these high-intensity riders, the lack of a cumulative goal, combined with high initial fatigue, results in a negative day^2 interaction. Consequently, while daily targets are effective for immediate re-activation, they are less resilient to the mid-week ”burnout gradient” than their cumulative counterparts.

7.3. Categorical Temporal Mapping

While the quadratic models provide a useful overview of the ”U-shaped” effort curve, they risk smoothing over discrete daily behavioral shifts. To isolate the precise day-by-day impact of incentive introduction, we re-estimated the MixedLM using a categorical treatment of time. By dummy-coding each day and setting the pre-incentive period (Day -1) as the reference baseline, we isolated the ”Pure Shock” of the incentive from natural weekly fluctuations.

$$Orders_{it} = \beta_0 + \sum_{k \neq -1} \beta_k D_{it}^k + \gamma Incentive_{it} + \sum_{k \neq -1} \delta_k (D_{it}^k \times Incentive_{it}) + u_i + \epsilon_{it}$$

Where D_{it}^k represents a categorical indicator for the k^{th} relative day of the incentive cycle. This specification allows for an unconstrained estimation of the incentive’s impact at every discrete temporal point.

Table 24 Categorical Temporal Mapping: Interaction Effects

Variable	Coef.	Std.Err.	z	P> $\|z\|$	[0.025	0.975]
Intercept (Day -1, No Incentive)	20.089	0.209	95.962	0.000	19.678	20.499
Main Day Effects						
Day 1	-0.148	0.200	-0.739	0.460	-0.541	0.245
Day 2	0.207	0.201	1.031	0.303	-0.186	0.600
Day 3	-0.797	0.200	-3.982	0.000	-1.189	-0.404
Day 4	-0.425	0.200	-2.124	0.034	-0.817	-0.033
Day 5	-2.710	0.223	-12.172	0.000	-3.147	-2.274
Day 6	-1.252	0.234	-5.355	0.000	-1.710	-0.794
Main Incentive Effects						
Daily Hours	-6.207	0.237	-26.172	0.000	-6.672	-5.742
Daily Orders	-0.926	0.271	-3.418	0.001	-1.457	-0.395
Streak Hours	-2.013	0.238	-8.460	0.000	-2.479	-1.547
Streak Orders	-0.308	0.237	-1.303	0.193	-0.772	0.156
Interaction: Day 1 Shock Factor						
Day 1 × Daily Hours	4.287	0.300	14.291	0.000	3.699	4.875
Day 1 × Daily Orders	-0.230	0.338	-0.681	0.496	-0.892	0.432
Day 1 × Streak Hours	1.600	0.299	5.353	0.000	1.014	2.186
Day 1 × Streak Orders	-0.947	0.292	-3.241	0.001	-1.519	-0.374
Group Var (driver_id)	107.430	0.201				

Table 25 Categorical Temporal Dynamics: Interaction of Incentive Type and Relative Day

Interaction Term	Coef.	Std.Err.	z	P> $\|z\|$	[0.025	0.975]
Daily Hours						
Day 1	4.287	0.300	14.291	0.000	3.699	4.875
Day 2	4.027	0.301	13.401	0.000	3.438	4.616
Day 3	3.113	0.299	10.426	0.000	2.528	3.698
Day 4	1.872	0.297	6.301	0.000	1.290	2.455
Day 5	3.895	0.313	12.451	0.000	3.282	4.508
Day 6	2.067	0.327	6.313	0.000	1.426	2.709
Daily Orders						
Day 1	-0.230	0.338	-0.681	0.496	-0.892	0.432
Day 2	0.210	0.339	0.620	0.535	-0.454	0.875
Day 3	-0.579	0.338	-1.712	0.087	-1.242	0.084
Day 4	-0.659	0.339	-1.948	0.051	-1.323	0.004
Day 5	1.141	0.352	3.243	0.001	0.451	1.830
Day 6	0.556	0.367	1.513	0.130	-0.164	1.276
Streak Hours						
Day 1	1.600	0.299	5.353	0.000	1.014	2.186
Day 2	1.825	0.300	6.089	0.000	1.238	2.412
Day 3	0.697	0.298	2.339	0.019	0.113	1.281
Day 4	0.465	0.298	1.561	0.119	-0.119	1.049
Day 5	2.508	0.314	7.998	0.000	1.893	3.122
Day 6	1.758	0.330	5.320	0.000	1.110	2.406
Streak Orders						
Day 1	-0.947	0.292	-3.241	0.001	-1.519	-0.374
Day 2	-0.968	0.293	-3.307	0.001	-1.542	-0.394
Day 3	-1.741	0.292	-5.967	0.000	-2.313	-1.169
Day 4	-1.977	0.292	-6.771	0.000	-2.549	-1.405
Day 5	0.359	0.308	1.166	0.244	-0.244	0.962
Day 6	0.548	0.326	1.682	0.093	-0.091	1.187

Our transition to a categorical modeling framework reveals that the ”Shock Factor” is not merely a uniform reaction to financial rewards, but a sensitive behavioral response dictated by the specific architecture of the incentive. The most aggressive reaction was observed in the Daily Hours group, which functions as a high-arousal ”sprint” mechanism. For these drivers, the categorical interaction for Day 1 shows a surge of +4.28 orders, which is nearly eight times the natural Day 1 increase seen in the control group. However, this intensity is short-lived; our mapping confirms a rapid ”temporal resignation,” where the marginal boost drops by over 50% to just +1.87 by Day 4. This provides empirical evidence that while time-based daily targets are good at triggering immediate labor supply, they are simultaneously most susceptible to a steep burnout gradient.

On the other hand, Daily Orders performs very similarly to our baseline, having little significant changes compared to baseline.

In contrast, Streak Hours incentives can be characterized by strategic energy conservation rather than immediate exhaustion. While the initial Day 1 shock for this group is more moderate (+1.60), the day-by-day analysis reveals a ”recovery phase” later in the cycle. Unlike memory-less daily models, streak-based structures leverage a persistence effect. As the week progresses, these drivers demonstrate a secondary surge on Day 5 (+2.50), which effectively neutralizes the severe natural weekend slump (-2.71) observed in the baseline group. By creating a back-loaded motivation structure, streak-based incentives successfully transform a period of typical resignation into a high-engagement sprint toward a cumulative goal.

However, there is the negative shock associated with Streak Orders. Drivers in this category exhibited a significant performance dip on Day 1 (-0.94, p¡0.001) relative to their pre-incentive baseline. This suggests a sophisticated ”pacing strategy” unique to volume-based cumulative rewards. Rather than succumbing to initial novelty, these drivers appear to intentionally suppress effort on the first day of the streak. By preserving their physiological and psychological capital during the early stages, they prepare for the high-intensity requirement of the final days, illustrating that drivers are not just reactive participants but strategic actors who optimize their labor supply across the entire temporal architecture of the incentive.

8. Sequencing Analysis

To evaluate the behavioral impact of incentive sequencing, we employed a comparative time-series analysis. This was achieved by holding the current week’s incentive (I_t) constant while varying the preceding week’s incentive (I_{t-1}). We monitored three primary performance dimensions: Order Completion Volume, Active Labor Hours, and Acceptance Rate. These metrics served as proxies for overall driver productivity and engagement.

Although data were initially segmented by driver tier, the stability of behavior across tiers led to the use of a single, cumulative trend line for final reporting. All metrics were normalized to Day 1 of the previous week. This baseline approach isolates the ”incentive sequence” effect, providing a clear visualization of whether productivity increases or decays when transitioning between specific incentive structures.

8.1. Orders Completed (Separated by Driver Tiers)

Overall, for Orders Completed, all driver tiers tend to exhibit the same behavior pattern at their respective levels. However, the incentive sequence that had the most impact on total orders completed was Streak Orders followed by No Incentive, especially for the Top tier group of drivers. We

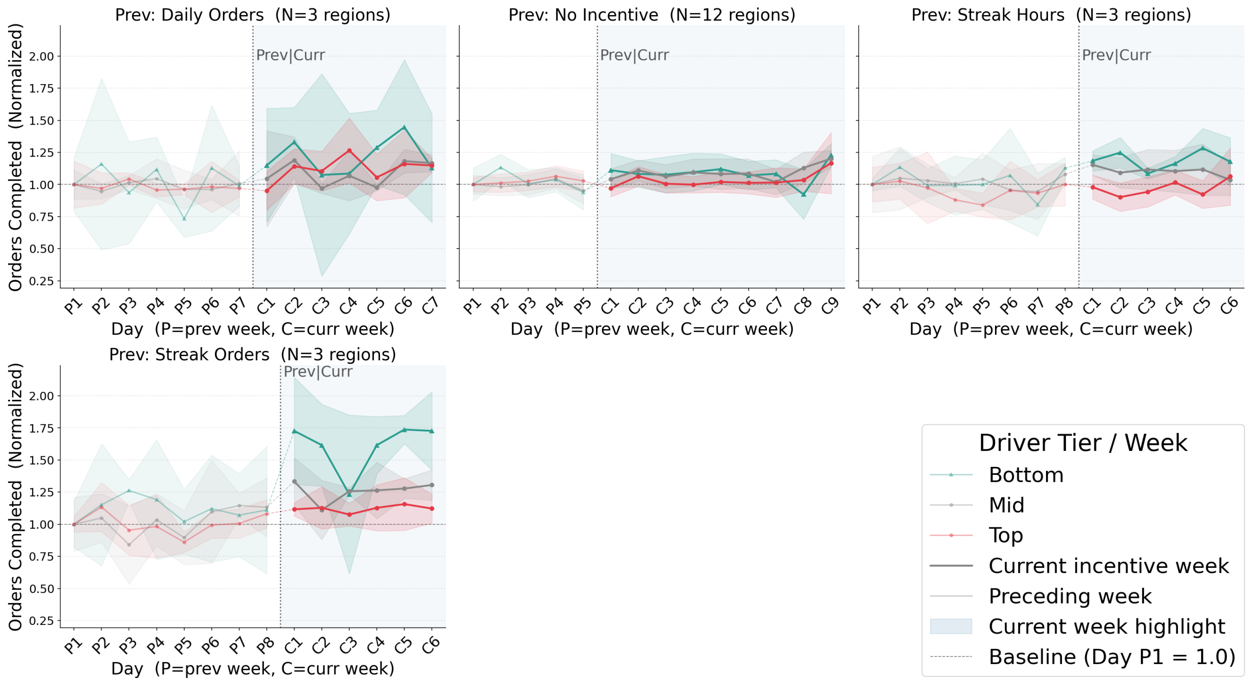


Figure 4 Orders completed holding Daily Hours Constant

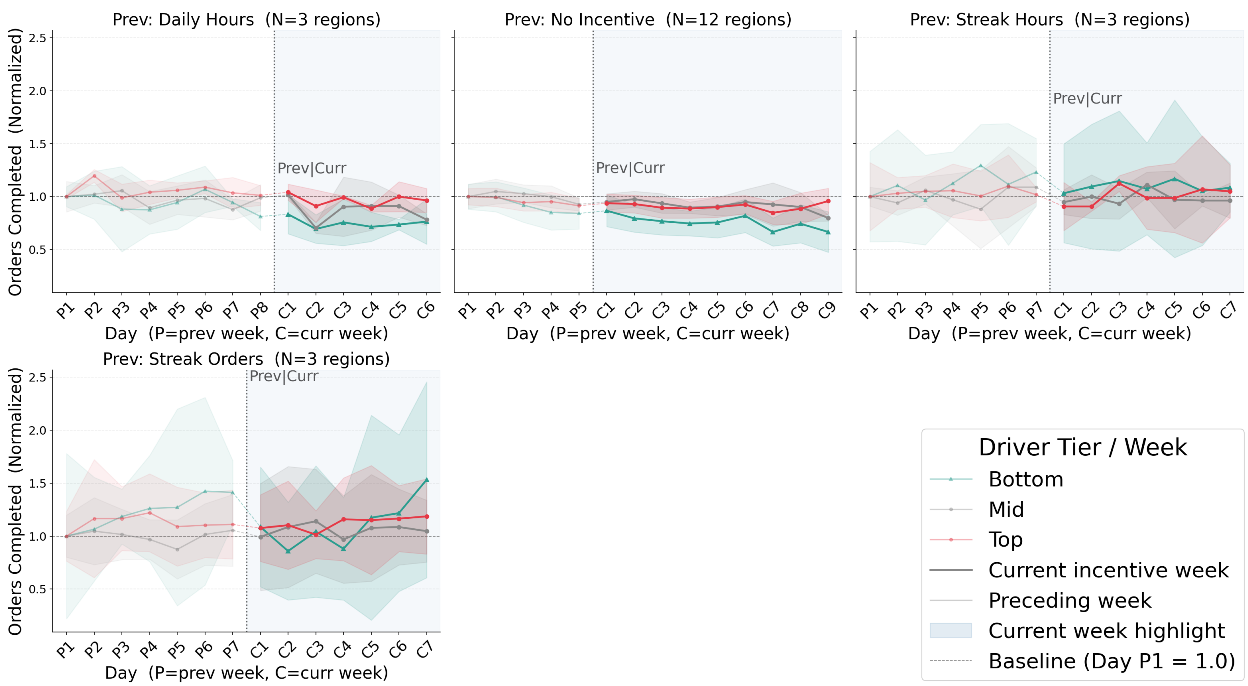


Figure 5 Orders completed holding Daily Orders Constant

can see that starting from the previous week, the total number of orders completed trends upward, with this behavior following into the next week with No Incentive, indicating that Streak incentive types might encourage the same type of behavior into the following week, even if the incentive is

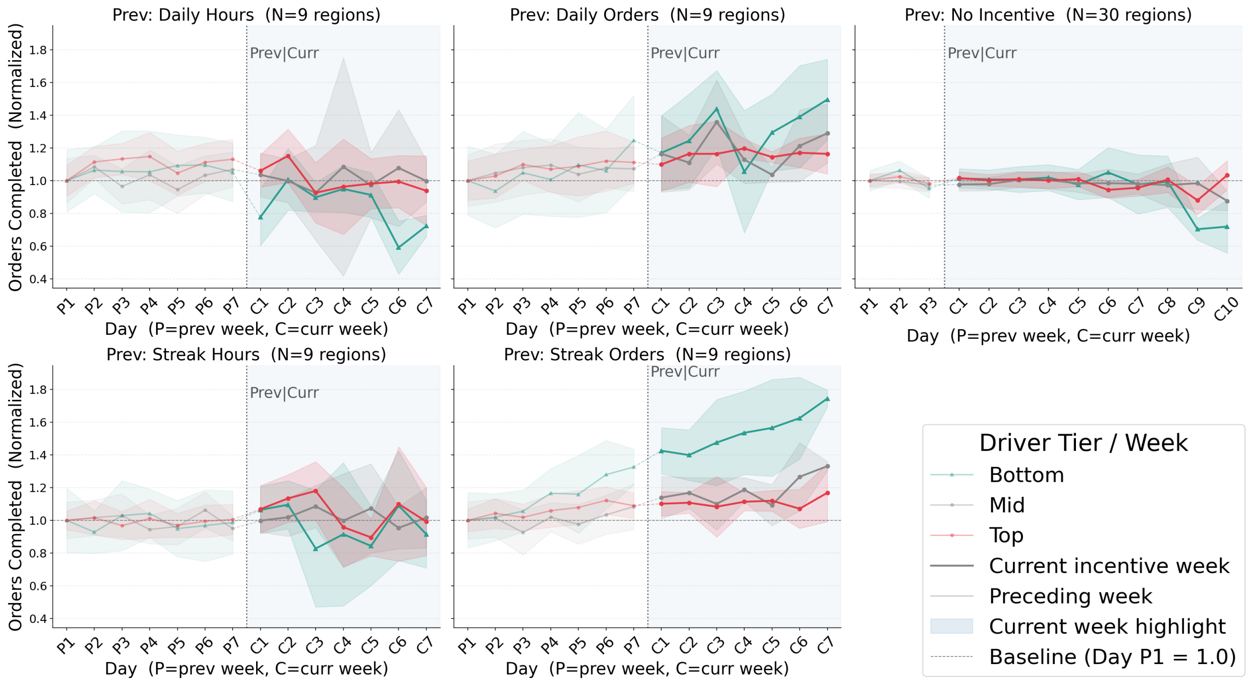


Figure 6 Holding No Incentive constant

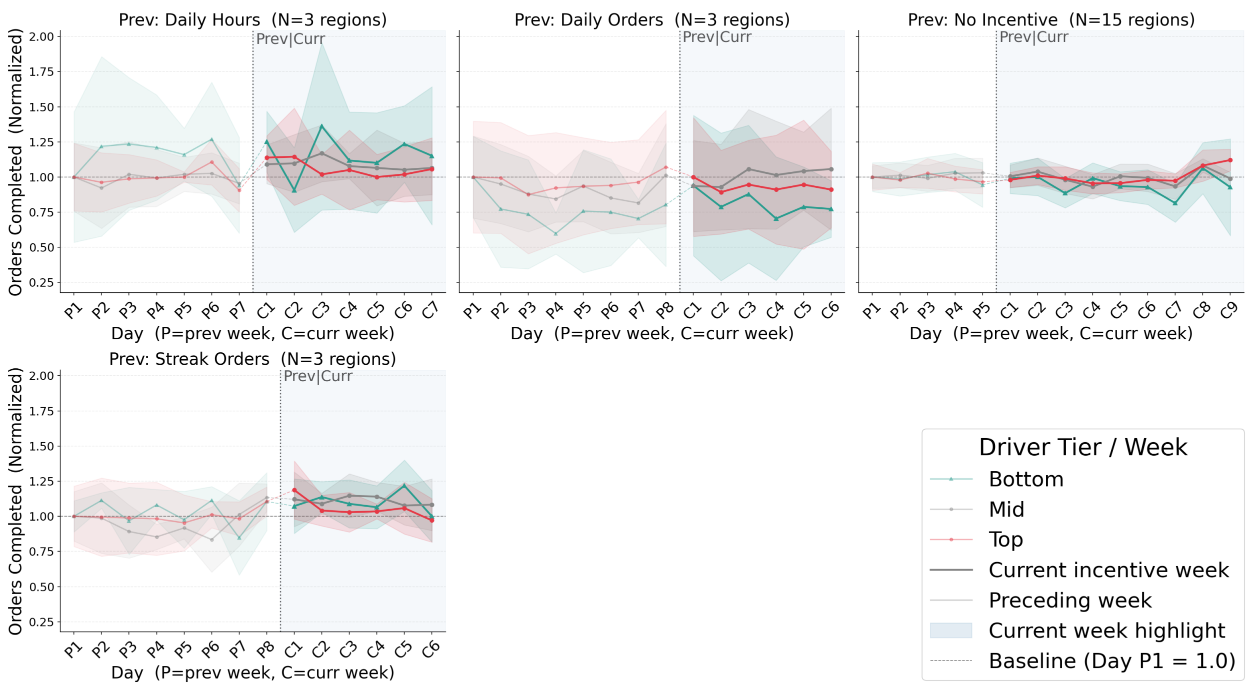


Figure 7 Orders completed holding Streak Hours constant

no longer offered (Figure 3). Throughout all the incentive types, if no incentive was offered the previous week, a similar behavior continues the next week, even if an incentive is offered.

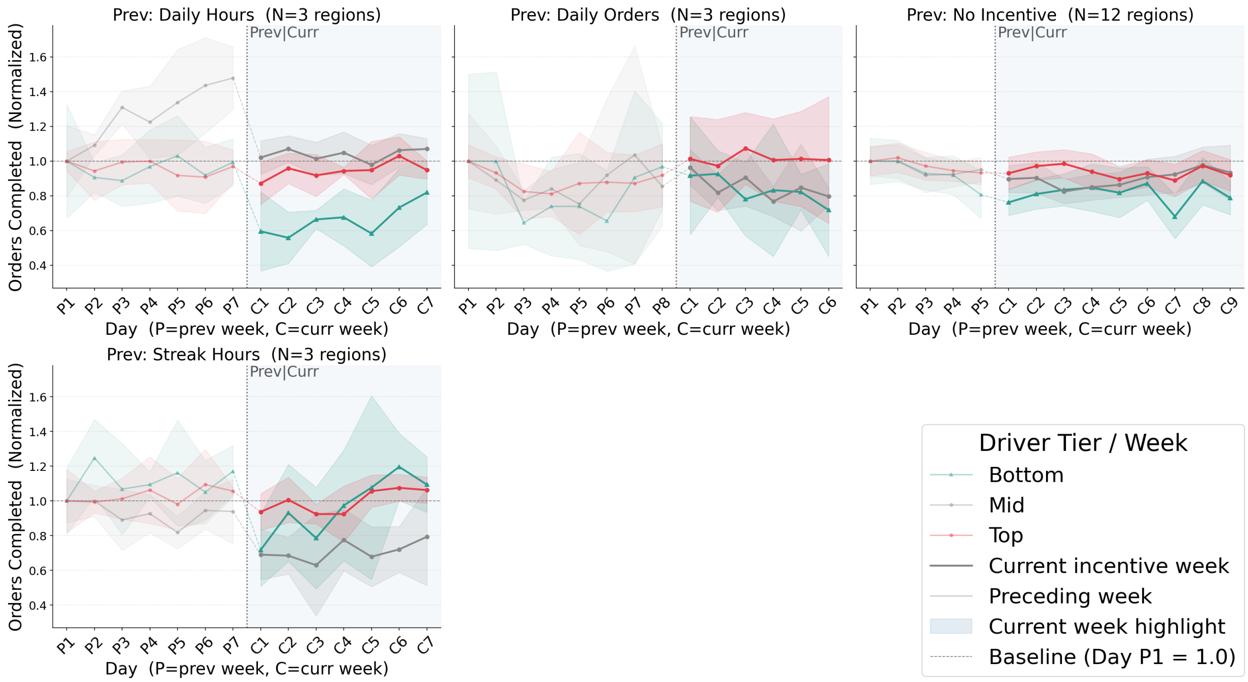


Figure 8 **Orders Completed holding Streak Orders constant**

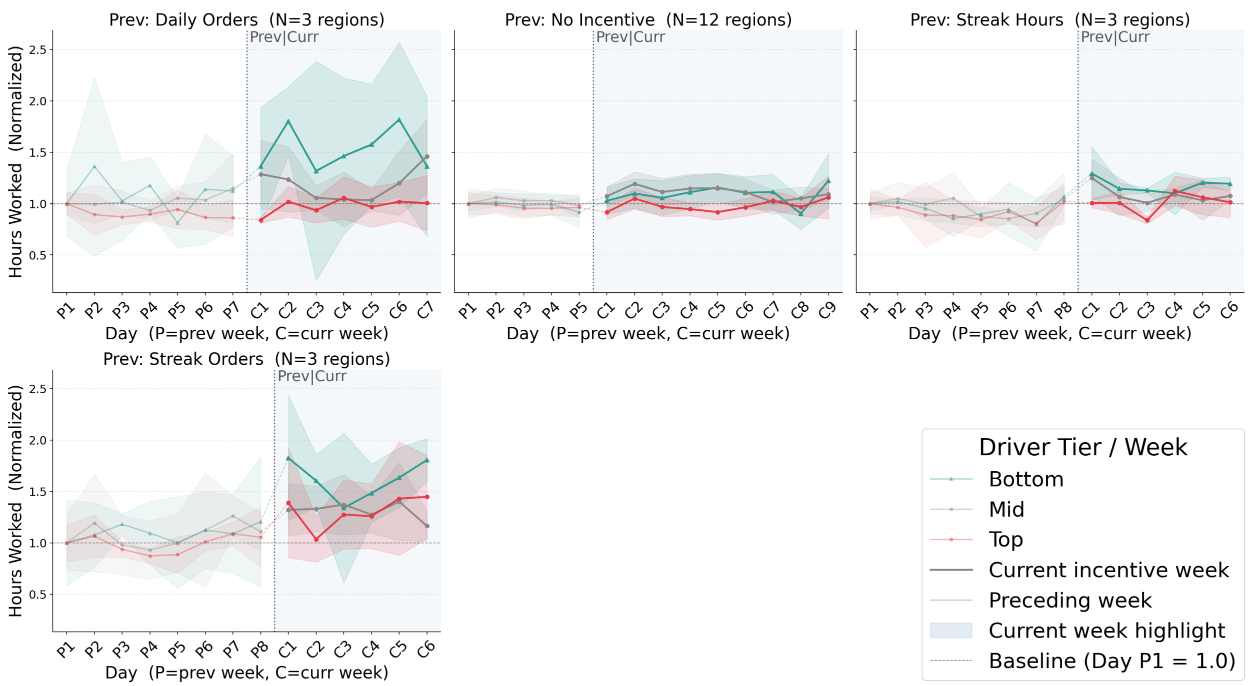


Figure 9 **Hours Worked holding Daily Hours constant**

8.2. Hours Worked (Separated by Driver Tiers)

Any incentive type preceding No Incentive exhibits the same behavior in terms of hours worked throughout the week for all driver types (Figure 8). All driver types in general seem to exhibit the

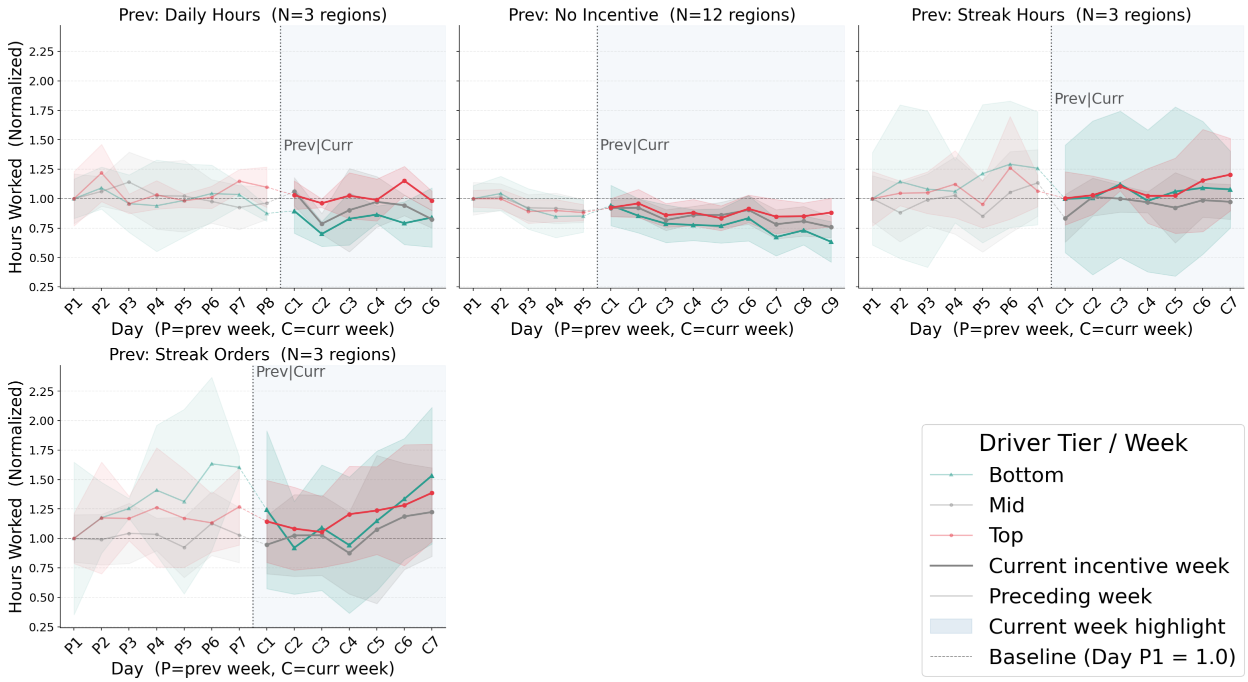


Figure 10 Hours Worked holding Daily Orders constant

Figure 11 Hours Worked holding No Incentive constant

same behavior in regards to each of their tiers in response to all incentive sequences. Since there was an upward trend in terms of orders completed from streak hours to no incentive, but not much

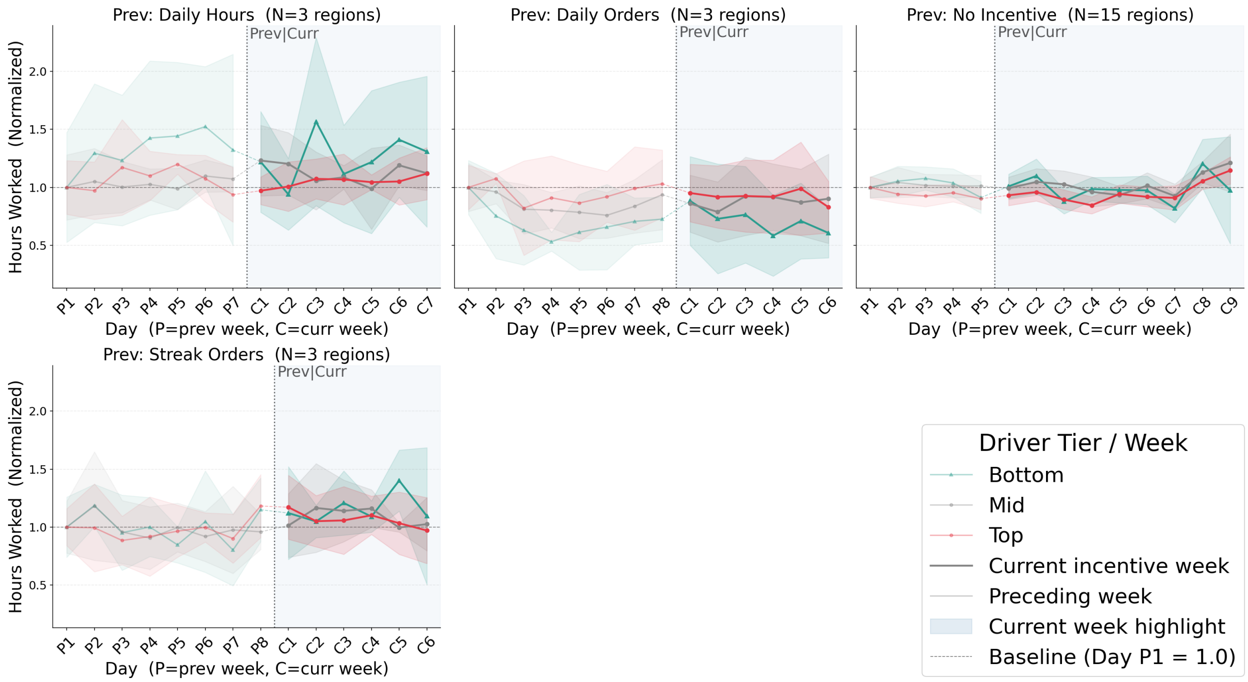


Figure 12 **Hours Worked holding Streak Hours constant**

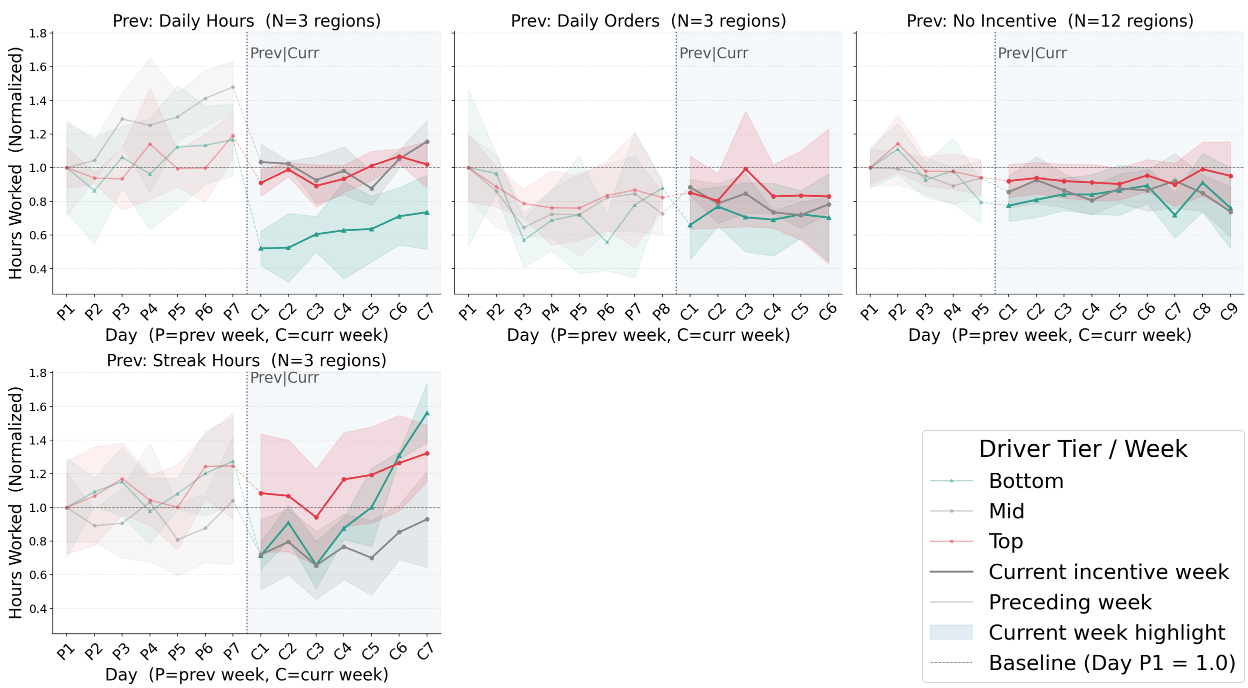


Figure 13 **Hours Worked holding Streak Orders constant**

of an upward trend in terms of hours worked for this same sequence, this indicates that drivers were being more productive, completing more orders per hour with this incentive sequence.

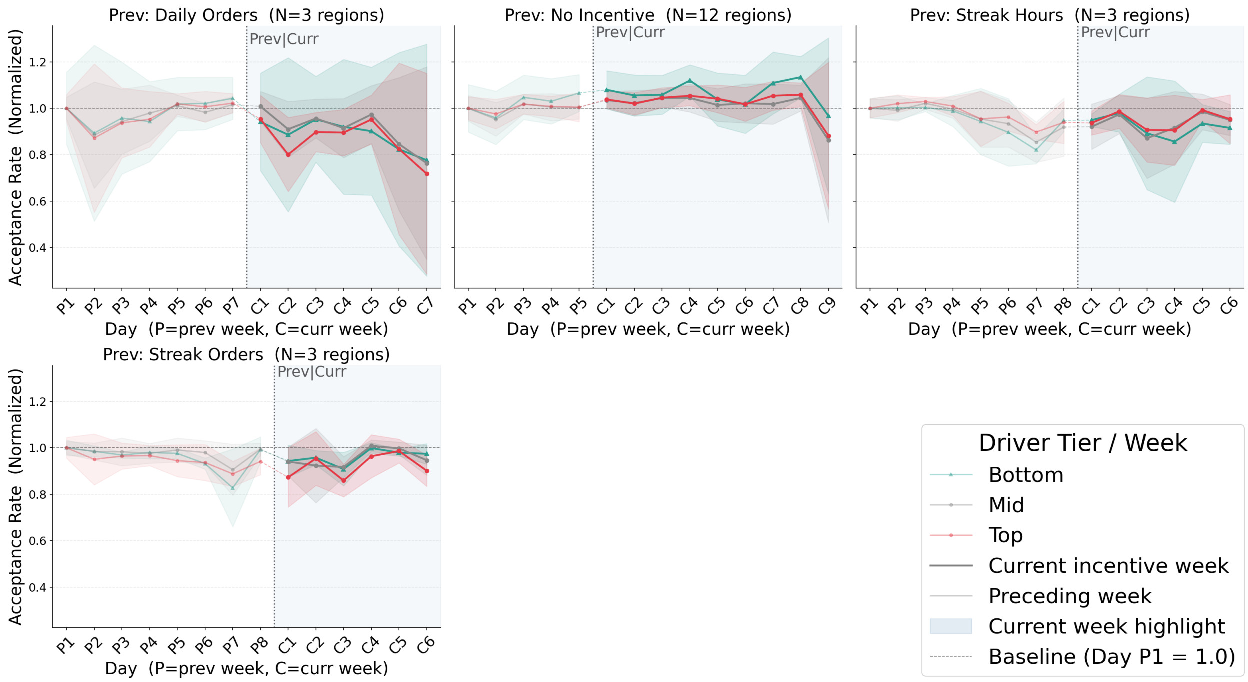


Figure 14 **Acceptance Rate holding Daily Hours constant**

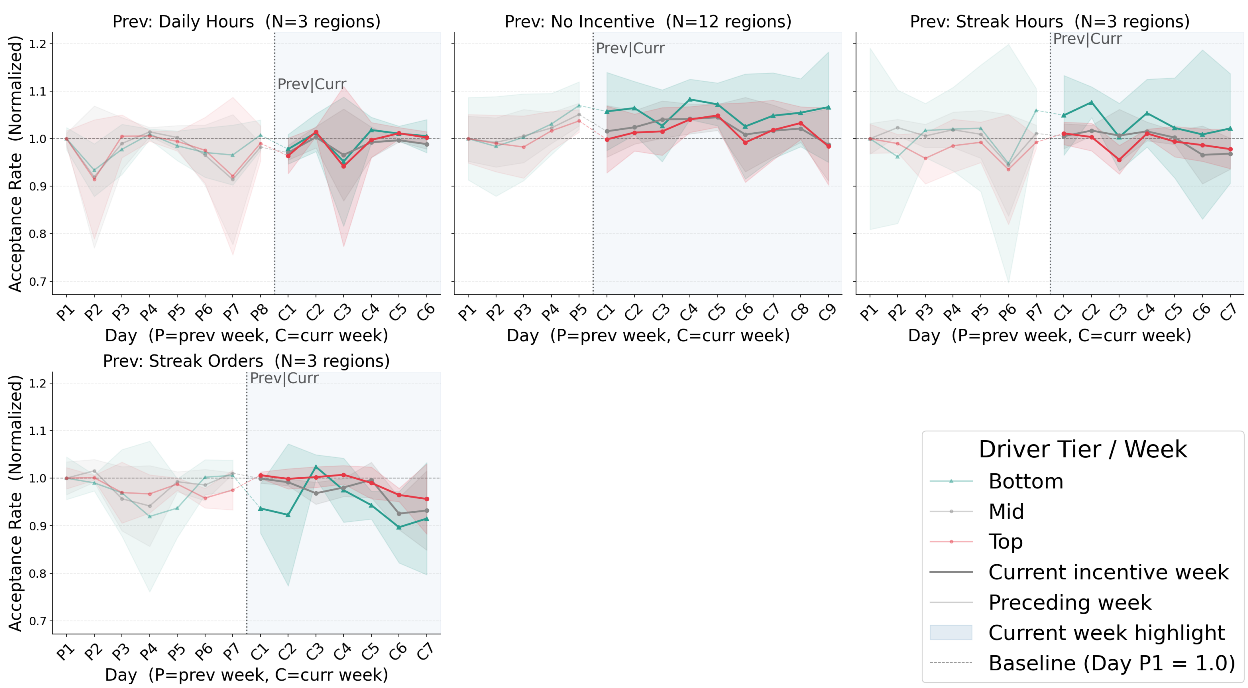


Figure 15 **Acceptance Rate holding Daily Orders constant**

8.3. Acceptance Rate (Separated by Driver Tiers)

There was a downward trend in acceptance rate from Streak Orders to No Incentive, indicating that this week may have had a higher quantity of orders that were submitted. However, drivers still had

Figure 16 **Acceptance Rate holding No Incentive constant**

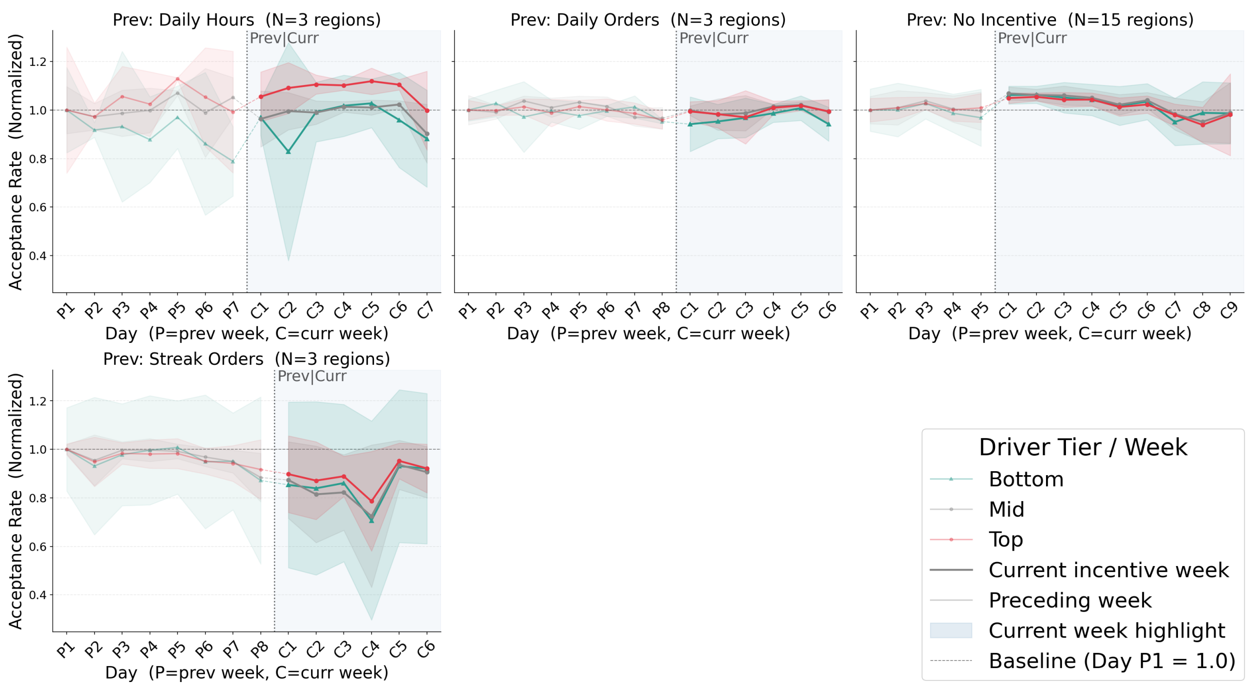


Figure 17 **Acceptance Rate holding Streak Hours constant**

a certain capacity, thus leading to a lower acceptance rate. Since this coincides with the upward trend of orders completed, but constant hours worked, the number of total orders may have had an affect on the behavior of drivers, that may not be solely because of the incentive sequencing. All

Figure 18 **Acceptance Rate holding Streak Orders constant**

drivers tend to exhibit the same behavior across all incentive types, consistent with the behavior seen with orders completed and hours worked. However, the upward trend of only the top group of riders indicate that the incentive sequencing may have been a factor in deciding to complete more orders with the total number of orders increasing.

8.4. Orders Completed (Trend Lines)

The behavior here is consistent with the ones displayed in the graphs separated by driver tiers. We combined all regions into one trend line for easy readability, because all drivers tended to display similar behaviors and reactions to each incentive sequence. With the trend line, the effect of incentive sequencing is made more apparent. The trend line was generated by taking the mean across all regions, rather than separating by rider tiers, and then taking the mean. From the trend line, we can see that Streak Orders than Daily Hours had an apparent increase from one week to another, with the second week having consistently more orders completed than the previous week (Figure 16). In addition, both Daily Orders than No Incentive, and Streak Orders than No incentive sequences have a slight trend upwards from one week to the next, indicating that Order-based incentives may be more effective than hour based ones for increasing the number of orders completed per day and maintaining this behavior into following weeks, even when the incentive is no longer active.

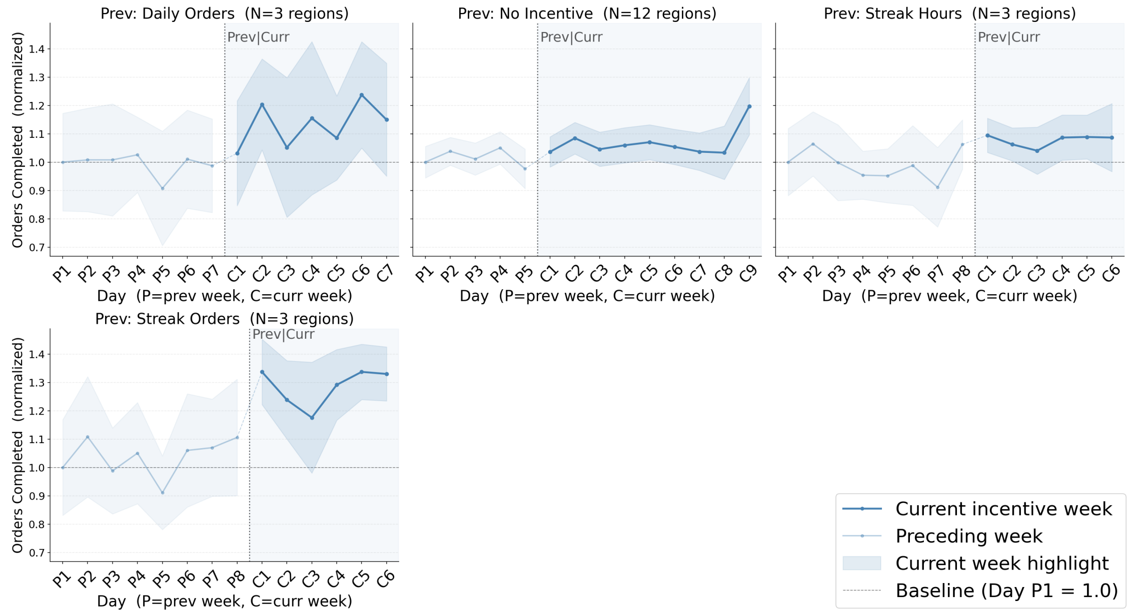


Figure 19 Orders Completed holding Daily Hours constant - Trend Line

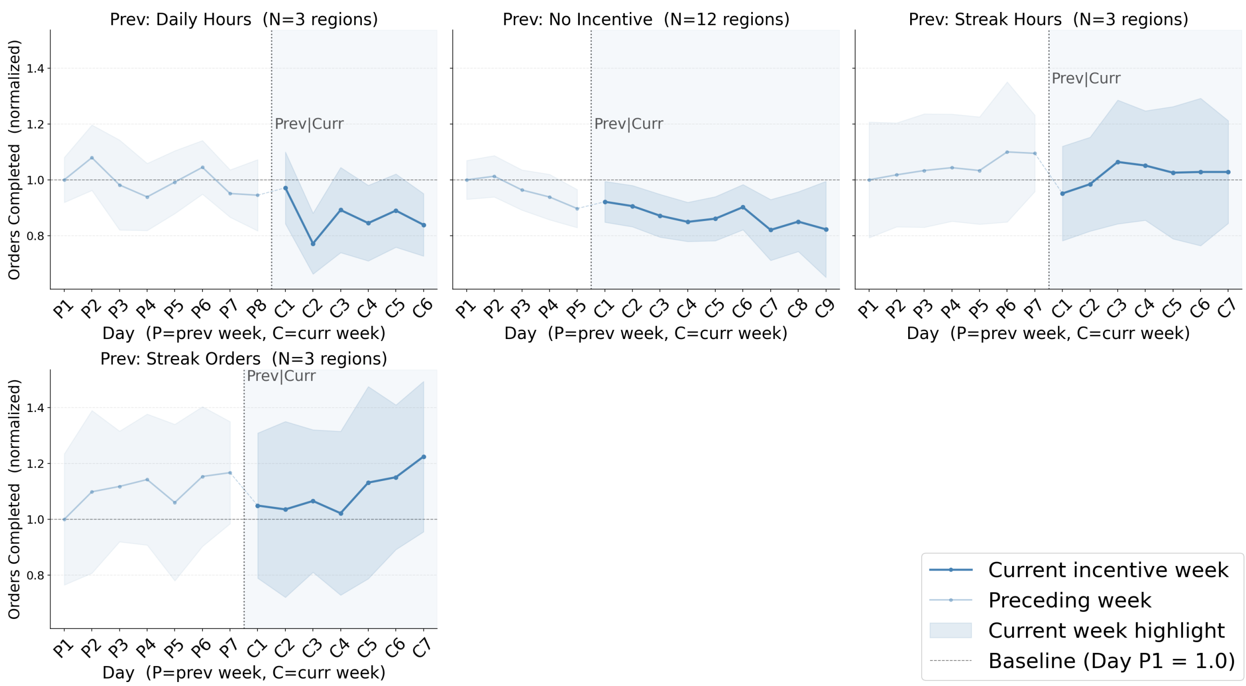


Figure 20 Orders Completed holding Daily Orders constant - Trend Line

8.5. Hours Worked (Trend Lines)

The behavior here is consistent with the ones displayed in the graphs separated by driver tiers. Trend lines were generated by taking the average across all 30 regions and all driver tiers. From these graphs, we can see a similar upwards trend in hours worked in the sequences from Daily Orders

Figure 21 Orders Completed holding No Incentive constant - Trend Line

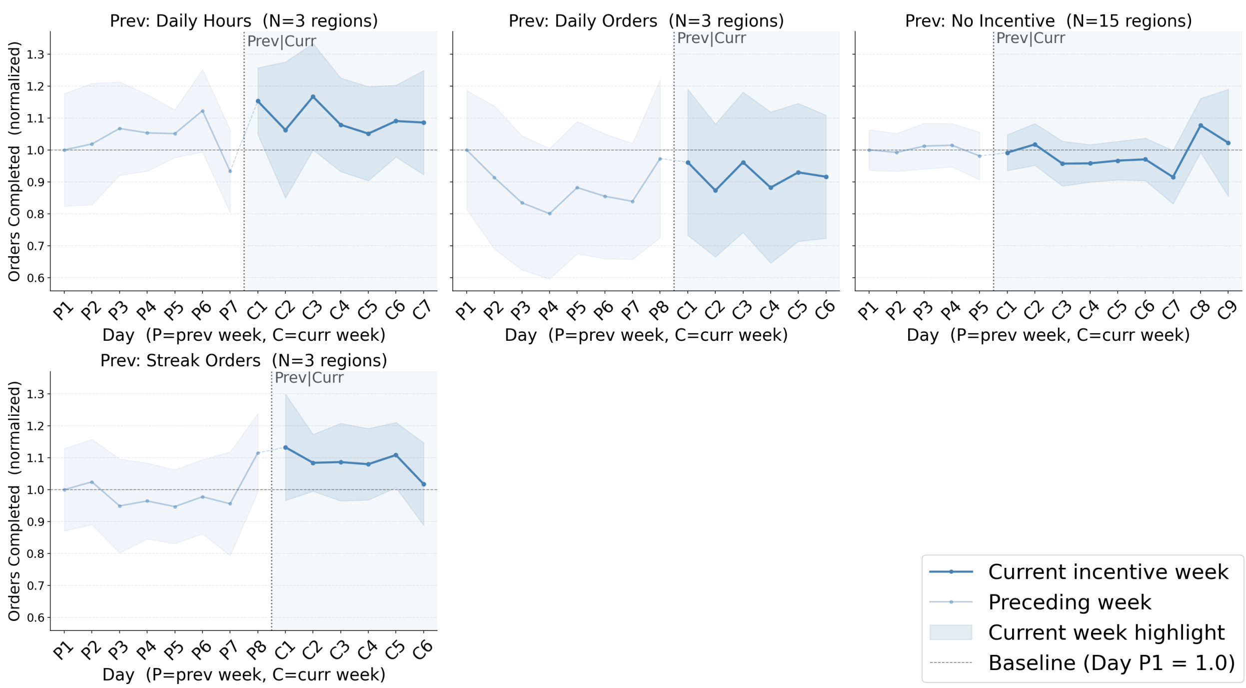


Figure 22 Orders Completed holding Streak Hours constant - Trend Line

to No Incentive, and from Streak Orders to No Incentive, however the increase is not as dramatic. This indicates that drivers are more productive with these incentive sequences.

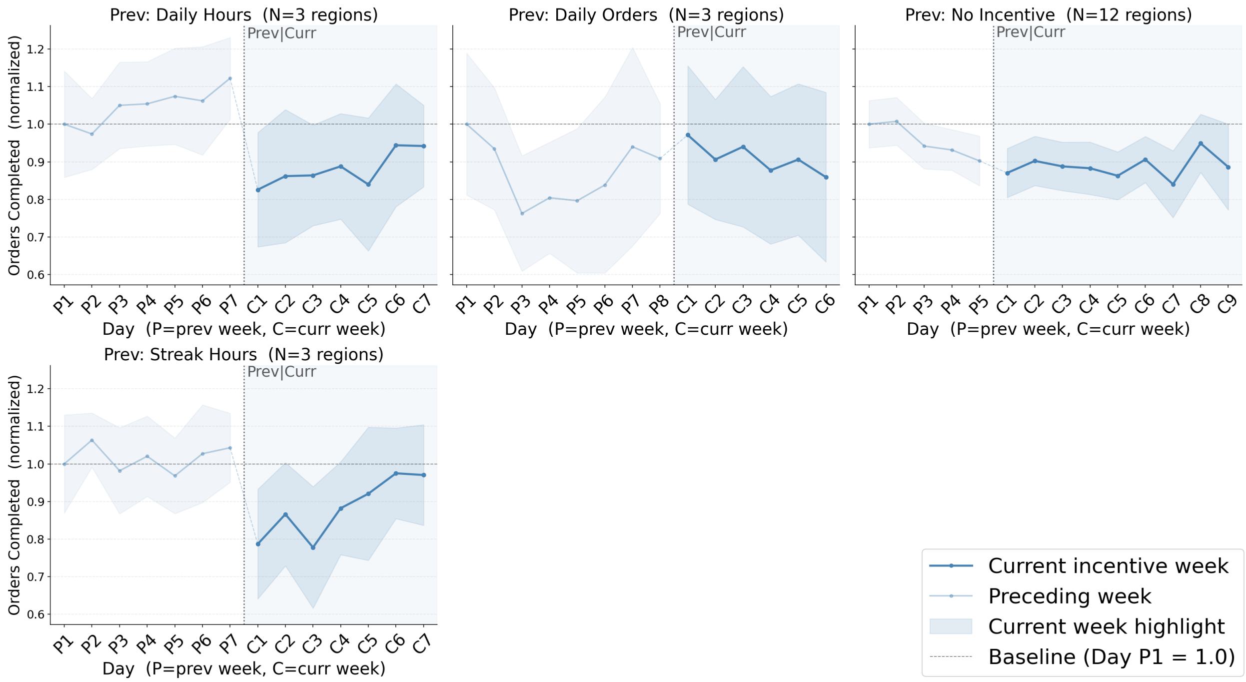


Figure 23 **Orders Completed holding Streak Orders constant - Trend Line**

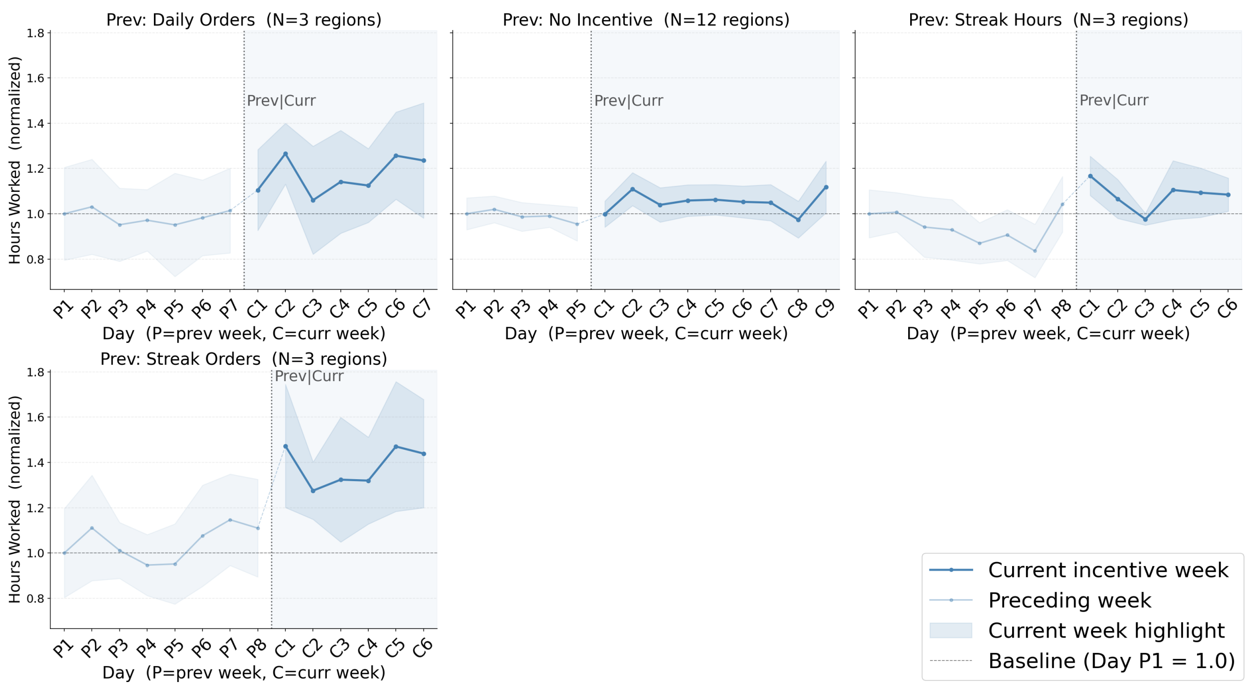


Figure 24 **Hours Worked holding Daily Hours constant - Trend Line**

8.6. Acceptance Rate (Trend Lines)

The behavior here is consistent with the ones displayed in the graphs separated by driver tiers. The trend lines were generated from the data across all 30 regions. From these graphs, we can see

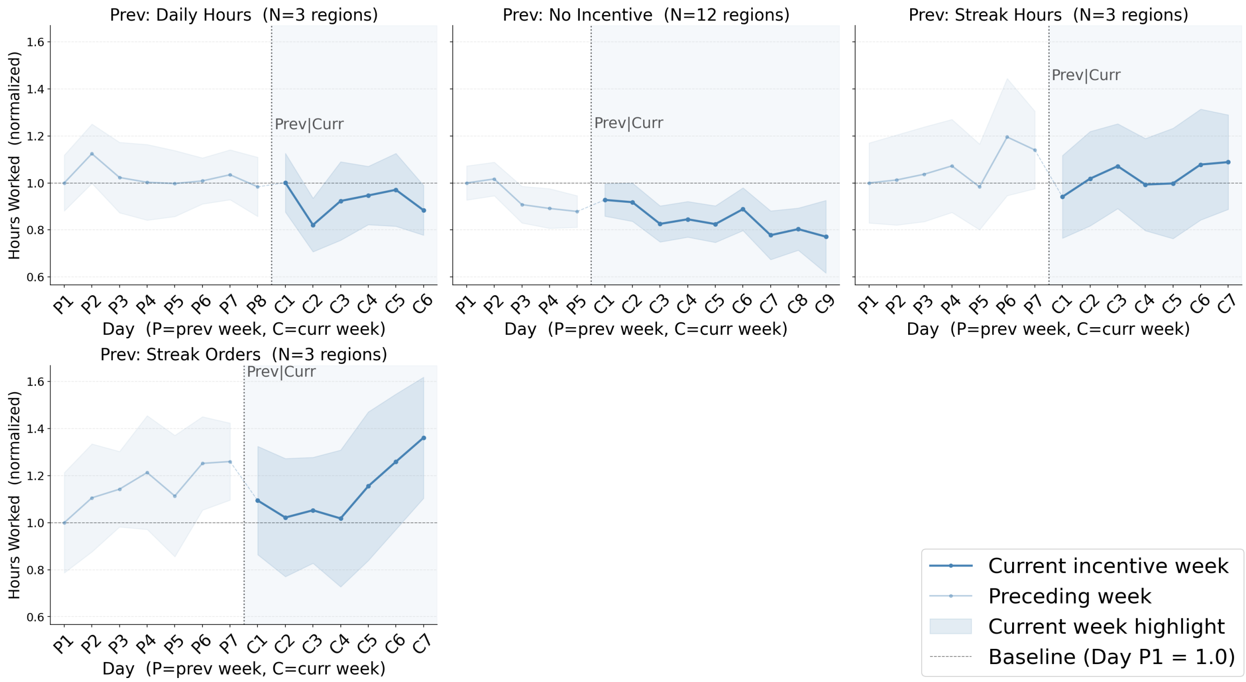


Figure 25 **Hours Worked holding Daily Orders constant - Trend Line**

Figure 26 **Hours Worked holding No Incentive constant - Trend Line**

that there is a very slight downward trend of acceptance rate from Daily Orders to No Incentive, indicating a true increase in productivity across all drivers with this sequence. This indicates that this sequence may have had a bigger role in the increased productivity of drivers, rather than an

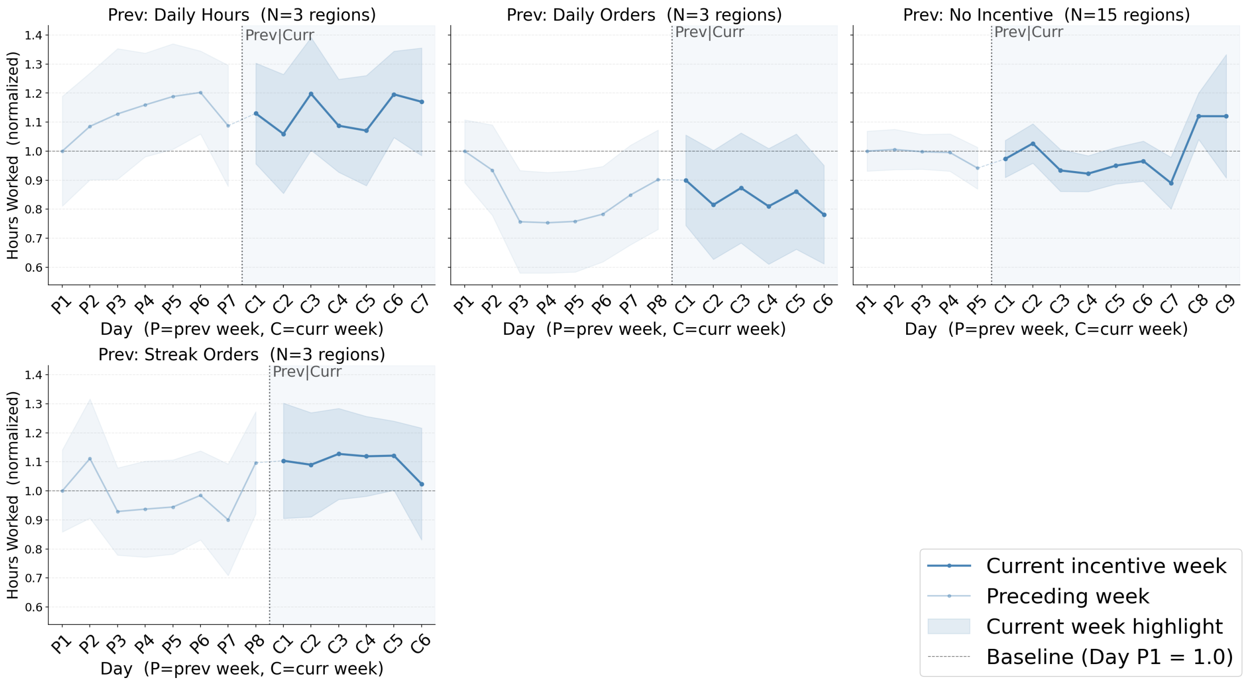


Figure 27 Hours Worked holding Streak Hours constant - Trend Line

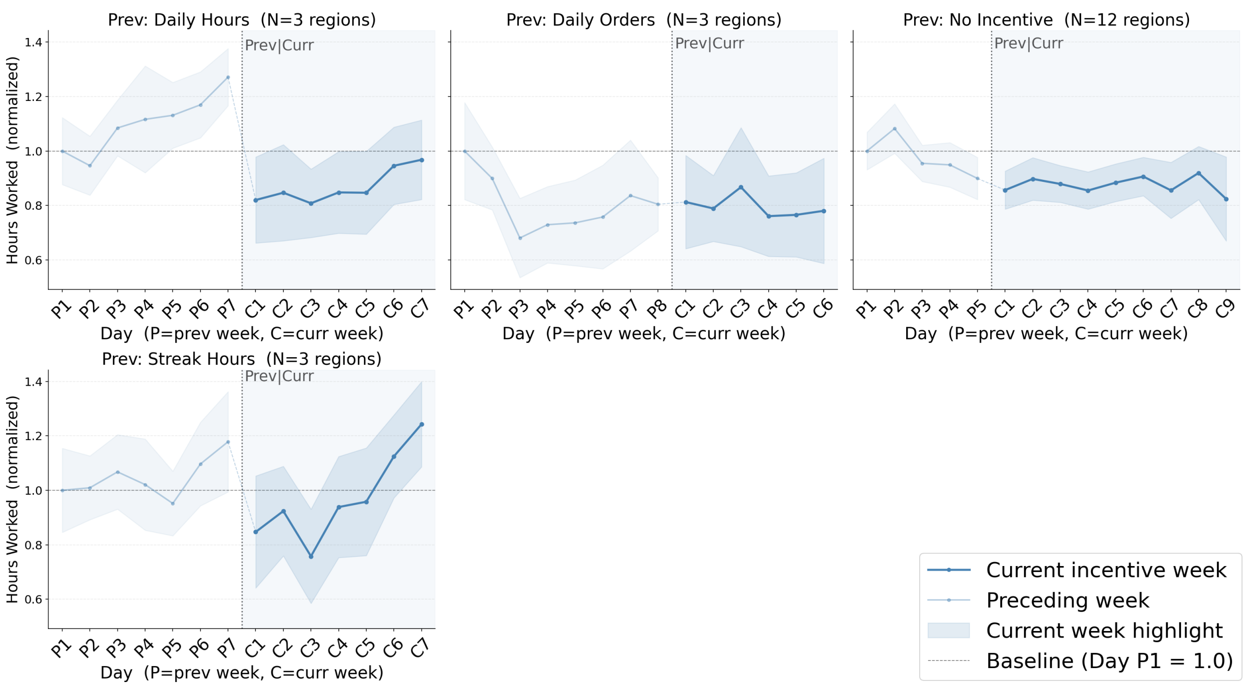


Figure 28 Hours Worked holding Streak Orders constant - Trend Line

increase number of orders playing a large role. However, from streak orders to no incentive, there is a very obvious decline in acceptance rate, indicating that there may have been an incline in the number of orders submitted that week, playing a bigger role in why the number of hours worked and

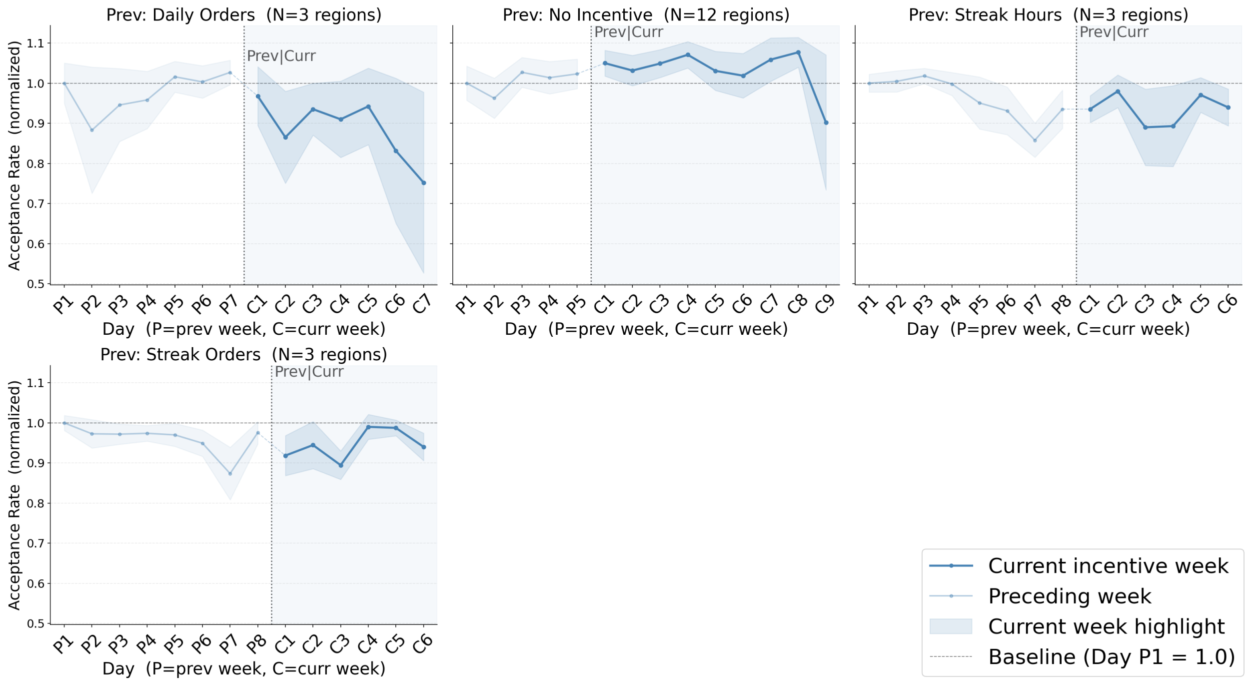


Figure 29 **Acceptance Rate holding Daily Hours constant - Trend Line**

Figure 30 **Acceptance Rate holding Daily Orders constant - Trend Line**

orders completed spiked with this incentive sequence. Across all incentive sequences, maintaining no incentive weeks maintains the most consistent acceptance rate.

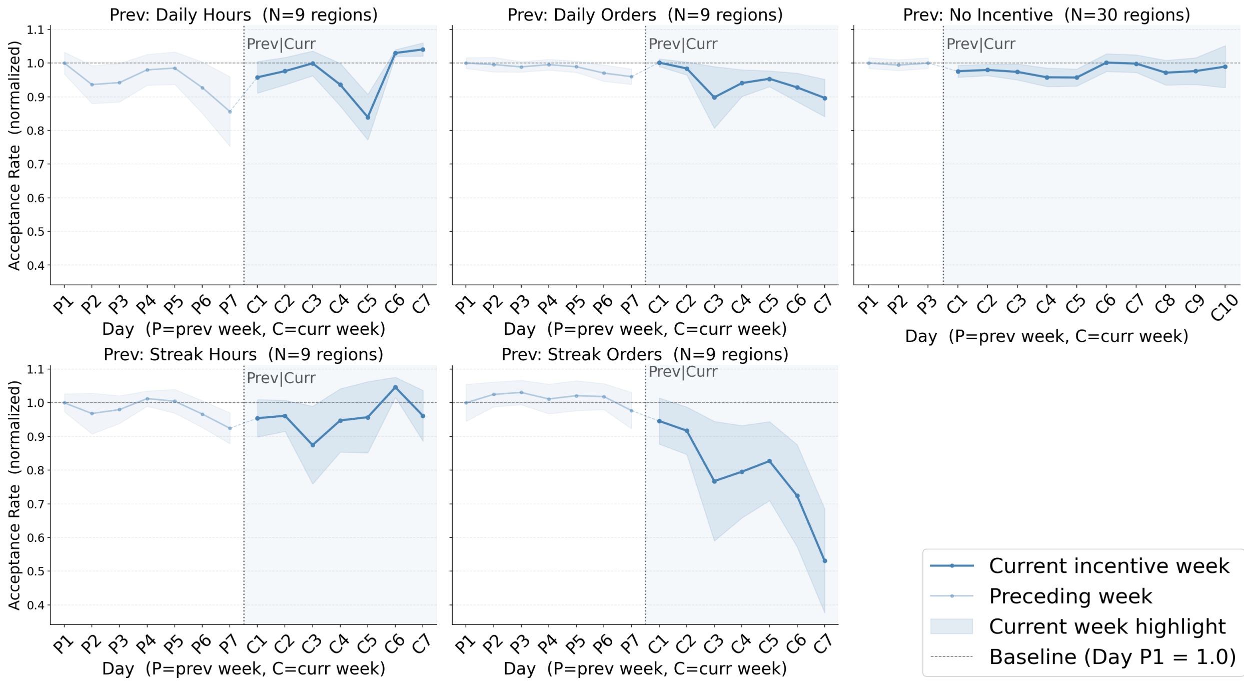


Figure 31 Acceptance Rate holding No Incentive constant - Trend Line

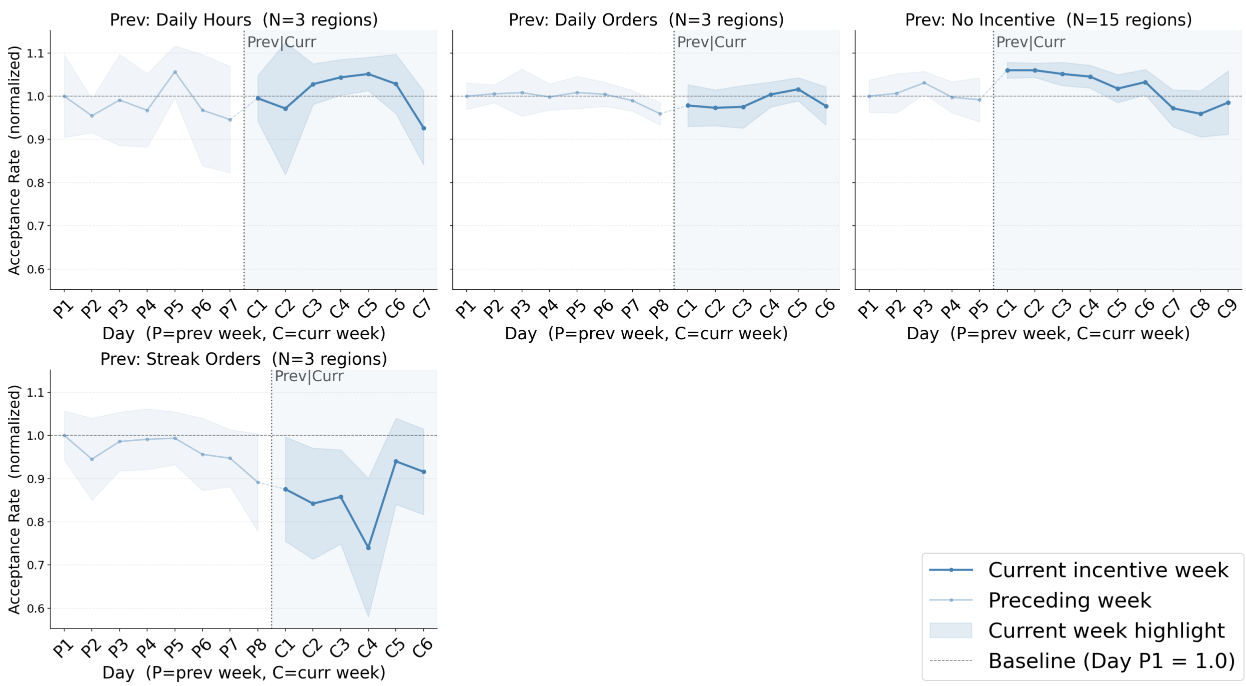


Figure 32 Acceptance Rate holding Streak Hours constant - Trend Line

8.7. T-test between different incentive sequences

To evaluate whether the ordering of incentives had a statistically significant impact on driver behavior, we conducted a series of pairwise Welch two-sample t-tests across all possible incentive

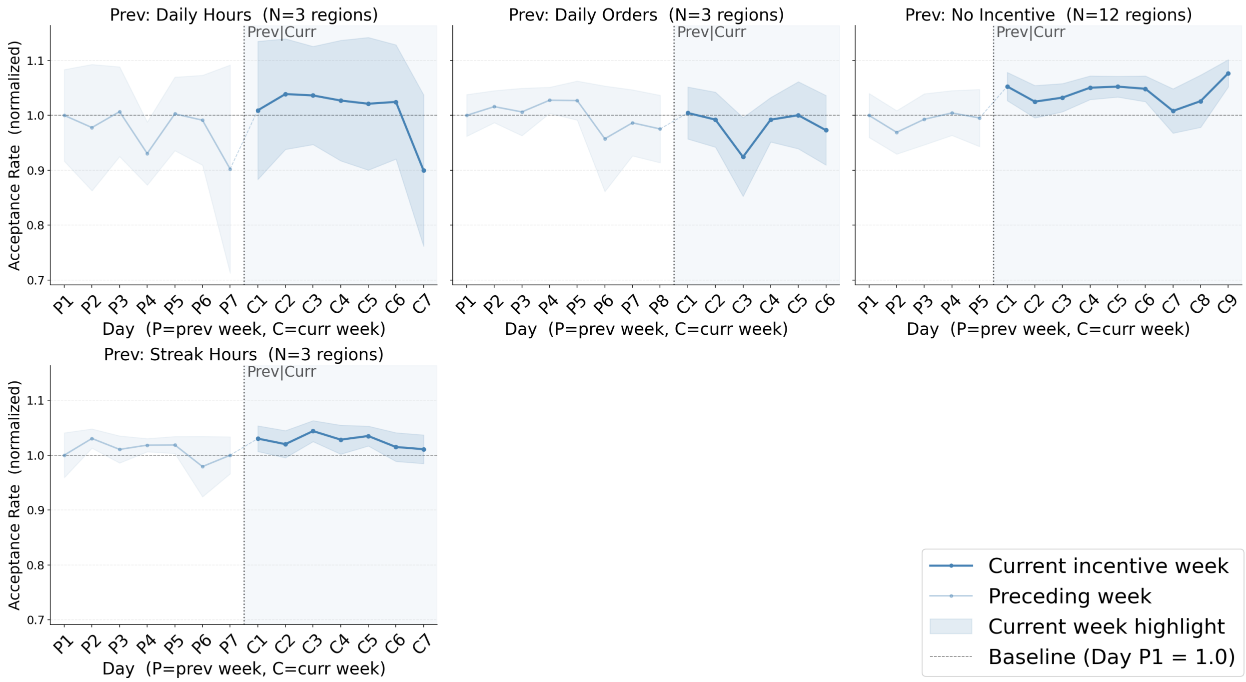


Figure 33 Acceptance Rate holding Streak Orders constant - Trend Line

sequence combinations observed in the data. The unit of analysis was a driver-week, defined as the mean outcome for a given driver within a given week under a specific incentive. This aggregation ensures that each observation is independent and not inflated by the number of orders a driver completed within that week.

An incentive sequence was defined as a pair of consecutive incentives assigned to the same driver within the same region across two adjacent weeks, denoted as ($\text{Incentive}_{t\text{-}1} \rightarrow \text{Incentive}_t$). For example, the sequence NI $\rightarrow$ DH indicates that a driver was assigned No Incentive in week t-1 and Daily Hours in week t. Only transitions within the same region were considered valid sequences, as incentives were assigned at the region level and cross-region transitions would not reflect a meaningful experimental condition.

For each unique predecessor incentive ($\text{Incentive}_{t\text{-}1}$), we identified all sequences sharing that same predecessor and compared the mean outcome during the successor week (Incentive_t) across every pair of such sequences. For instance, holding the predecessor constant at NI, we compared the mean acceptance rate of drivers in the NI $\rightarrow$ DH group against those in the NI $\rightarrow$ SH group. This design ensures that any observed difference in the successor week outcome is attributable to the successor incentive itself, since the prior week's incentive is identical across both groups being compared.

The outcome variable was measured as the driver's mean acceptance rate, orders completed, or hours worked during the successor week, depending on the specification. Welch's t-test was used in place of Student's t-test as it does not assume equal variances between groups, which is appropriate given the heterogeneous nature of driver behavior across regions and incentive types. To account for the inflated risk of false positives arising from multiple comparisons, a Bonferroni correction was applied by multiplying each raw p-value by the total number of comparisons conducted. Only sequences with a minimum of 30 observations were included to ensure sufficient statistical power. Results were considered statistically significant at the 5% level after Bonferroni correction.

Table 26 Pairwise T-Test Results for Incentive Sequences - Orders Completed

seq_A	seq_B	mean_A	mean_B	t_stat	p_value	p_bonferroni	sig_bonferroni
NI → DH	NI → SH	22.9227	19.6983	7.7618	0.0000	0.0000	True
NI → SH	NI → DO	19.6983	22.4930	-6.0493	0.0000	0.0000	True
NI → SH	NI → SO	19.6983	21.9769	-5.8715	0.0000	0.0000	True
DO → NI	DO → SO	24.2563	19.2793	5.3970	0.0000	0.0000	True
DO → NI	DO → DH	24.2563	19.7591	5.3656	0.0000	0.0000	True
SH → NI	SH → SO	24.3480	20.4789	5.1661	0.0000	0.0000	True
SO → SH	SO → DH	24.2590	21.1490	4.7041	0.0000	0.0001	True
SH → NI	SH → DH	24.3480	21.0965	4.0941	0.0000	0.0014	True
SO → NI	SO → DH	23.3756	21.1490	3.5235	0.0004	0.0133	True
DH → SO	DH → NI	25.5200	22.5730	3.5017	0.0005	0.0147	True

Notes: Table only shows significant sequence comparisons (e.g from the table, we can see that Streak Orders to Streak Hours is statistically significantly better than Streak orders to Daily Orders in terms of orders completed); SO = Streak Orders, SH = Streak Hours, DH = Daily Hours, DO = Daily Orders, NI = No Incentive. Standard errors computed using Welch two-sample t-test.

Table 27 **Pairwise T-Test Results for Incentive Sequences - Acceptance Rate**

seq_A	seq_B	mean_A	mean_B	t_stat	p_value	p_bonferroni	sig_bonferroni
SO → NI	SO → DO	0.7818	0.9339	-17.6703	0.0000	0.0000	True
SO → SH	SO → DO	0.7940	0.9339	-16.5594	0.0000	0.0000	True
SO → NI	SO → DH	0.7818	0.9163	-15.6951	0.0000	0.0000	True
SO → SH	SO → DH	0.7940	0.9163	-14.5433	0.0000	0.0000	True
DH → SH	DH → DO	0.8515	0.9353	-13.2642	0.0000	0.0000	True
SH → NI	SH → DO	0.8066	0.9321	-10.0974	0.0000	0.0000	True
NI → SH	NI → SO	0.9034	0.8522	9.6927	0.0000	0.0000	True
DO → DH	DO → SH	0.7866	0.9255	-9.9826	0.0000	0.0000	True
SH → NI	SH → SO	0.8066	0.9187	-9.9230	0.0000	0.0000	True
DH → NI	DH → DO	0.8582	0.9353	-9.4710	0.0000	0.0000	True
DO → DH	DO → SO	0.7866	0.9226	-8.9876	0.0000	0.0000	True
DO → NI	DO → DH	0.9054	0.7866	8.9299	0.0000	0.0000	True
NI → DH	NI → SH	0.8612	0.9034	-8.0360	0.0000	0.0000	True
SH → DH	SH → SO	0.8778	0.9187	-7.0983	0.0000	0.0000	True
SH → DH	SH → DO	0.8778	0.9321	-7.0076	0.0000	0.0000	True
DH → SO	DH → DO	0.8805	0.9353	-6.4253	0.0000	0.0000	True
NI → SO	NI → DO	0.8522	0.8878	-6.0814	0.0000	0.0000	True
SH → NI	SH → DH	0.8066	0.8778	-6.0361	0.0000	0.0000	True
NI → DH	NI → DO	0.8612	0.8878	-4.5664	0.0000	0.0002	True
DH → SO	DH → SH	0.8805	0.8515	3.3370	0.0009	0.0267	True

Table 28 Pairwise T-Test Results for Incentive Sequences - Hours Worked

seq_A	seq_B	mean_A	mean_B	t_stat	p_value	p_bonferroni	sig_bonferroni
DO → NI	DO → SO	5.3430	3.6542	9.4570	0.0000	0.0000	True
NI → DH	NI → SH	4.8591	4.1229	8.7960	0.0000	0.0000	True
NI → SH	NI → SO	4.1229	4.7437	-7.9484	0.0000	0.0000	True
NI → SH	NI → DO	4.1229	4.8596	-7.8982	0.0000	0.0000	True
SH → NI	SH → DO	5.2623	4.1181	6.2413	0.0000	0.0000	True
SH → NI	SH → DH	5.2623	4.3402	5.7144	0.0000	0.0000	True
DO → NI	DO → DH	5.3430	4.3133	5.6639	0.0000	0.0000	True
SH → NI	SH → SO	5.2623	4.4325	5.4337	0.0000	0.0000	True
DO → SH	DO → SO	4.6437	3.6542	5.0286	0.0000	0.0000	True
SO → SH	SO → DH	5.2623	4.6570	4.2008	0.0000	0.0009	True
DH → SO	DH → NI	5.3070	4.6558	3.9501	0.0001	0.0026	True
DO → NI	DO → SH	5.3430	4.6437	3.8819	0.0001	0.0034	True
SO → NI	SO → DH	5.1462	4.6570	3.5460	0.0004	0.0123	True
DO → DH	DO → SO	4.3133	3.6542	3.3234	0.0009	0.0283	True
DH → SO	DH → SH	5.3070	4.7909	3.3224	0.0009	0.0283	True

Overall, across all three metrics (Hours Worked, Orders Completed, and Acceptance Rate), transitioning to No Incentive after an active incentive tends to produce stronger performance than transitioning from one active incentive to another. This suggests that even when an incentive is removed, the increased productivity that occurs as a result of a prior week's incentive carries over into the following week, highlighting behavior persistence. One possible explanation is that drivers are anticipating the return of the incentive, or that repeated exposure to incentive structures leads to habit formations.

Out of all Orders Completed transitions, Daily Hours tends to outperform Streak Hours as a successor incentive when coming from No Incentive (22.9 vs. 19.7 orders on average). One possible explanation is that the requirements for Daily Hours are less difficult to reach, only requiring a daily goal rather than multiple cumulative goals throughout the week, making the target more attainable and thus more motivating. Overall, the largest mean order counts are observed among drivers transitioning into No Incentive (e.g., DO→NI yields 24.3 orders vs. DO→SO yields 19.3 orders), a pattern that repeats consistently for SH→NI and SO→NI as well.

Acceptance Rate yields the most significant incentive pair comparisons across all three outcomes. Daily Orders consistently produces the highest acceptance rates as a successor incentive, regardless of what came before it. For example, SO→DO yields 0.934, SH→DO yields 0.932, and DH→DO yields 0.935, all significantly higher than nearly every alternative successor. On the other hand, Streak Orders tends to produce low acceptance rates as a successor (around 0.78–0.79). We hypothesize that streak-based goals may seem unattainable or intimidating to drivers, making them less likely to engage compared to Daily goals, which only require reaching a daily threshold to receive a benefit rather than sustaining performance across multiple checkpoints throughout the week.

For Hours Worked, transitioning into No Incentive produces the highest mean hours across all successor comparisons (e.g., DO→NI yields 5.34 hours, SH→NI yields 5.26 hours, SO→NI yields 5.15 hours), all significantly outperforming transitions into active incentives. This further supports the behavior persistence finding observed across the other metrics, suggesting that productivity gains from prior incentive exposure are not immediately lost when incentives are withdrawn.

8.8. Overall Incentive Sequencing Findings

Table 29: Incentive Transition Effects

Previous	**%Imp.**	**Avg%**	**Med%**	**Qualitative Summary**
Current Incentive: Daily Hours				
Daily Orders	61.9	5.97	0.72	Modest broad gains; most metrics improve
Streak Orders	57.1	7.05	0.33	Strong avg gain but high variance; wages dip
No Incentive	47.6	0.09	-0.28	Near-zero net effect; slight downward drift
Streak Hours	42.9	2.93	-0.37	Mixed results; more decline than improvement
Current Incentive: Daily Orders				
Streak Orders	57.1	-0.24	1.21	Slight net loss but median positive; wages hold
Daily Hours	52.4	-2.88	0.34	Evenly split; high effect size, orders suffer
Streak Hours	42.9	-1.32	-0.45	Largest effect size; mostly negative transition

Continued on next page

Table 29: Incentive Transition Effects (continued)

Previous	**%Imp.**	**Avg%**	**Med%**	**Qualitative Summary**
No Incentive	28.6	-2.70	-1.51	Weakest transition; most metrics decline
Current Incentive: No Incentive				
Daily Orders	90.5	4.95	3.44	Near-universal improvement; strongest transition
Streak Orders	85.7	10.85	4.11	Highest avg gain overall; wages surge
Streak Hours	85.7	5.62	3.67	Broad consistent gains; acceptance rate lags
Daily Hours	52.4	0.05	0.42	Marginal net effect; roughly balanced outcomes
No Incentive	47.6	0.46	-0.02	Baseline: no large effects; slight improvement
Current Incentive: Streak Hours				
Streak Orders	76.2	5.50	3.33	Strong broad gains; high effect size
Daily Hours	76.2	4.40	2.63	Consistent positive transition; wages improve
Daily Orders	28.6	-1.00	-2.22	Most metrics decline; high effect size
No Incentive	14.3	-2.22	-1.61	Worst transition overall; near-universal decline
Current Incentive: Streak Orders				
Streak Hours	61.9	-2.50	0.55	Majority improve; negative avg driven by outliers
Daily Hours	47.6	-4.26	-1.67	Net negative; high effect size, hours suffer
Daily Orders	38.1	-1.06	-0.09	Mostly declining metrics; near-zero median
No Incentive	14.3	-4.25	-3.00	Worst for Streak Orders; steep broad decline

To assess the impact of incentive sequencing, we employed a comparative analysis across seven distinct performance metrics. These metrics were categorized into four operational domains: Labor Supply (orders completed and hours worked), Compensation (total wages), Driver Engagement (acceptance rates), and Logistical Efficiency (average time per order, drop distance, and basket size). By holding the current week's incentive constant and observing the variance from the preceding week's performance, we isolated the "carryover effect" of specific incentive transitions.

The values presented in Table 8 were derived by calculating a composite Improvement Index (%Imp.). This index represents the percentage of the seven tested metrics that demonstrated a favorable week-over-week trend following an incentive change. "Improvement" was defined directionally:

we tracked numerical increases for volume and engagement metrics, while tracking numerical decreases for logistics-based friction, such as delivery duration and restaurant-to-customer displacement. This standardized aggregation allows for a holistic comparison of how different historical incentive structures prime or inhibit driver productivity in subsequent periods.

References

Momentum Works (2024) Food delivery platforms in southeast asia 2024. URL https://momentum.asia/insights/detail/food-delivery-platforms-in-southeast-asia-2024.

Thairath Money (2024) Line man 44% grab 1 . URL https://www.thairath.co.th/money/tech_innovation/tech_companies/2813800.

E-Companion to *"Designing and Sequencing Incentives for Gig Economy Workers"*

Appendix A: Additional Detail